CANADIAN JOURNAL OF PSYCHOLOGY

VOL. 9, NO. 4, DEC. 1955

PUBLISHED FOR THE

CANADIAN PSYCHOLOGICAL ASSOCIATION BY THE

UNIVERSITY OF TORONTO PRESS

CANADIAN JOURNAL OF PSYCHOLOGY

VOLUME 9, NO. 4 DECEMBER, 1955

PUBLISHED FOR THE
CANADIAN PSYCHOLOGICAL ASSOCIATION BY THE
UNIVERSITY OF TORONTO PRESS

AUTHORIZED AS SECOND-CLASS MAIL, POST-OFFICE DEPARTMENT, OTTAWA

CANADIAN JOURNAL OF PSYCHOLOGY

THE CANADIAN JOURNAL OF PSYCHOLOGY is published quarterly in March, June, September, and December. *Annual subscription*, $4.00; *single number*, $1.00.

Subscriptions. Orders and correspondence regarding subscriptions, change of address, and purchase of back numbers should be sent to:

THE SECRETARY TREASURER, CANADIAN PSYCHOLOGICAL ASSOCIATION
3600 McTavish Street, Montreal 2, Que.

Contributions. Original manuscripts and correspondence on editorial matters should be sent to:

THE EDITOR, CANADIAN JOURNAL OF PSYCHOLOGY
100 St. George Street, Toronto 5, Ont.

Information for contributors. The *Journal* publishes experimental and theoretical articles in all recognized fields of psychology. Contributors need not be members of the CPA or residents of Canada. Manuscripts should be double-spaced throughout, and should follow standard practice as regards tables, references, etc. "Immediate publication" (i.e. in the next issue to go to press) can be arranged for authors willing to pay the extra costs involved.

CANADIAN PSYCHOLOGICAL ASSOCIATION, 1955–1956

1956 Annual Meeting
Ottawa, Ontario
June 7, 8 and 9

The Canadian Psychological Association also publishes *The Canadian Psychologist*, which is distributed to members only. *Editors*: JEAN GARNEAU and ERNEST POSER, Montreal.

INFLUENCE OF A STIMULUS VARIABLE ON STORIES TOLD TO CERTAIN TAT PICTURES[1]

P. LYNN NEWBIGGING
University of New Brunswick[2]

L. K. FRANK (1) defines a projective technique as follows:

> A projective method for the study of personality involves the presentation of a stimulus situation designed or chosen because it will mean to the subject not what the experimenter has arbitrarily decided it should mean . . . but rather whatever it must mean to the personality who gives it, or imposes upon it, his private, idiosyncratic meaning and organization.

This cannot be accepted as an adequate definition of the TAT and particularly of those pictures in the TAT series that depict people. Although Stein (3) speaks of "unstructured situations" in referring to TAT pictures, he goes on to describe common themes in stories told in response to them. The fact that common themes are readily identifiable must mean that the meaning and organization imposed upon the pictures are not completely idiosyncratic but, rather, that private meaning is imposed *within the limits set by the structure of the pictures themselves.* How otherwise are common themes to be accounted for?

One aspect of this structure is the depiction of persons of a certain age and in certain physical relation to other persons or objects; another aspect is the "feeling tone" of the picture as determined by the posture, facial expression, and other expressive behaviour of the persons depicted. It is this latter aspect of structure, and, specifically, its determining effect on the feeling tone of stories told to these pictures, that is investigated in this research.

The Problem

Recent research on the relationship between personality and perception has shown conclusively that an individual's experience, attitudes, needs, etc., give him a "set" which influences his perception. For the most part these researches have involved the use of stimuli made ambiguous in various ways (e.g. tachistoscopic exposure, reduced illumination), and have required of a group of subjects, homogeneous with respect to "set," a relatively simple response such as identification of the stimulus.

[1]Adapted from part of a thesis submitted in 1952 in fulfilment of the requirements for the degree of Doctor of Philosophy in the Department of Psychology, University College London. I wish to thank Professor R. W. Russell and Mr. J. W. Whitfield for their assistance with the work.

[2]Now at McMaster University, Hamilton, Ont.

The present research is similar, except that a series of stimuli or configurations ordered with respect to a specific criterion was used, and a creative response was required of the subjects.

Specifically, a series of 10 TAT pictures was ranked by a number of judges with respect to the happiness-sadness of the scene. Subjects were instructed to "make up as happy a story as you can" to each picture. The subjects were thus instructed, rather than simply "to tell a story," in order to approximate the conditions under which the TAT is ordinarily used. That is, in the usual clinical situation, the subject presumably has some dominant need(s) which he projects and which (also presumably) can be determined from an analysis of his stories. The results obtained were analysed to see if the stories varied in happiness with the rated happiness of the pictures, or whether the "set" created by the instructions had so affected the response that no such relationship obtained.

Method

Material

The experimenter and one other person selected from the complete series of TAT pictures 12 which they agreed could be placed in order from happy to sad. These 12 were then presented to each of 5 judges by the Paired Comparison Method. Each judge was asked to choose the happier picture of each pair presented. The average correlation[3] between the orders of the individual judges was $\tau = 0.64$ ($P < .01$). The 12 pictures selected, and the order in which they were placed on the basis of the total number of choices each received, are as follows: 8GF, 10, 2, 7GF, 7BM, 6GF, 1, 5, 4, 9GF, 6BM, 3GF. After an inspection of the data, pictures 2 and 4 were discarded from the series. These differed by only 2 choices and 1 choice, respectively, from adjacent pictures.

The order in which the pictures are listed in the preceding paragraph (with pictures 2 and 4 omitted) is the order in which they were presented to the subjects.

Subjects

Fifty-two students from University College acted as subjects, 26 being male and 26 female. Twelve university departments were represented. Age range was from 17 to 32 (to nearest birthday) and mean age was 21.

Procedure

Each subject was seated at a narrow table opposite to the experimenter and given the following instructions:

I am going to show you some pictures one at a time and I want you to make up as happy a story as you can about each picture. Tell what led up to the event shown in the picture, what is happening at the moment and then what the outcome is likely to be. Remember, tell as happy a story as you can. Allow 3 or 4 minutes for each story. They do not have to be elaborate just so long as they have a beginning, a middle, and an end.

---

[3]M. G. Kendall's (2) rank order correlation methods were used to compute all correlation coefficients reported.

I am going to write the stories down as you tell them, so if you do not go too quickly that will help. There are 10 pictures altogether; here is the first one.

The pictures were then presented one at a time.

The following was recorded:

1. Response time: this was the time elapsing, to the nearest second by stopwatch, between the presentation of the picture and the moment the subject began his story.

2. The story: each story was taken down in long hand, as nearly as possible word for word as the subject told it. The number of words in each story was counted and recorded.

During an additional session the 10 pictures were presented to each subject by the Paired Comparison Method.[4] The subject was instructed as follows:

Now I am going to show you these pictures two at a time and I want you to judge which is the happiest of each pair. If the one you judge to be the happiest is on your right just say "right"; if the happiest one is on your left just say "left." Do the same for all pairs.

The subjects were encouraged to choose between all pairs rather than to make judgments of equality.

Since it was impossible to know to what extent, if any, the choices made by this group were influenced by their having told stories about the pictures, it was considered desirable to have the latter judged by another group of subjects. This group was composed of 7 men and 4 women. Their ages ranged from 18 to 27 (to nearest birthday) with a mean of 22.5. The pictures were presented to this group exactly as to the larger group.

The 520 stories (10 stories told by each of 52 subjects) were then rated by 5 raters, 3 of whom were men and 2 women. Three were graduate students in psychology, two professional psychologists.

Ratings were made on a 7-point scale which had defined degrees of happiness for each point from +3 through 0 to −3. Each rater was supplied with a set of the 520 stories, each typewritten on a separate slip of paper, a sheet of instructions telling him how the ratings were to be made, and seven large sheets, one for each point on the scale. The stories were done up in two bundles, one containing 20 stories and the other containing 500. Each of the large sheets had the appropriate symbol at the top, i.e. +2, −1, etc., designating its point on the scale, and two examples of the type of story that was to be placed at that point.[5]

The instructions were to the effect that the rater was first to read the stories on the large sheets and then to spread these out in order in front of him. He was then to take the smaller bundle, read each story, rate it on the basis of his impression of it as a whole, and place it on that large sheet which defined the appropriate degree of happiness. After completing these 20 stories he was to collect them, rate the remaining 500 similarly, and then re-rate the first 20.

The bundles of stories given to each rater were randomized by shuffling for both pictures and subjects. Thus each rater rated the stories in a different order.

The 14 stories used to define the points on the scale were chosen by the experimenter and one other person from the entire 520 stories. The use of complete stories

[4]This session followed the storytelling one by a minimum of four days, and a number of other experiments intervened.

[5]A copy of the rating scale may be obtained by writing to the author.

to deline the points on the scale was decided on after several other methods had been judged unsatisfactory.

Results

Two types of statistical method were used for the analysis of results. These were (1) Kendall's rank order correlation methods for all measures of relationship, and (2) medians and associated measures of variability for comparisons of distributions. The reason for these choices was that all distributions were skewed.

Judgment of the Pictures

The medians and semi-interquartile ranges for the distributions of choices of pictures for both groups of subjects (of 52 and 11 members, respectively) are given in Table I.

TABLE I

Medians and Semi-interquartile Ranges of Distributions of Choices of Pictures by the Two Groups of Subjects

Picture	Group of 52 subjects		Group of 11 subjects	
	Mdn.	Q	Mdn.	Q
8GF	8.24	0.95	8.25	1.08
10	6.50	1.96	8.12	0.72
7GF	6.57	1.30	6.62	1.22
1	5.92	1.34	5.33	1.00
6GF	5.33	1.62	5.25	1.18
5	3.88	1.30	4.00	1.35
7BM	3.30	1.46	3.38	1.44
9GF	2.50	1.26	3.25	2.00
6BM	1.37	0.86	1.20	0.64
3GF	0.92	1.38	0.42	0.90

In Table I the pictures are arranged approximately in order of judged happiness; the larger the median, the happier the picture was judged to be. The order of magnitude of the medians for the two groups agrees exactly except for pictures 10 and 7GF, which are reversed in order. For the two groups, the correlation between the medians ranked in order of magnitude is $\tau = 0.96$ ($P < .01$).

The shapes of the distributions of choices for the 10 pictures are of interest and can best be shown graphically. Figures 1, 2, and 3 show the distributions of choices for pictures 8GF, 6GF, and 3GF, respectively, made by the group of 52 subjects. These figures exemplify the trend for the distributions of all 10 pictures to proceed in shape from a marked negative skew, in the case of picture 8GF, through a more normal form,

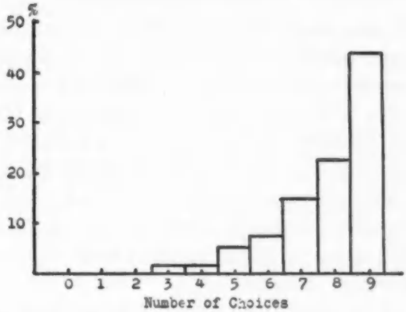

FIGURE 1. Percentage of subjects giving specified number of choices to picture 8GF.

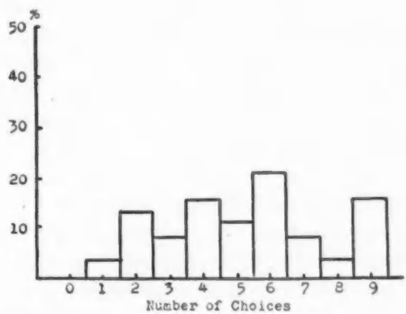

FIGURE 2. Percentage of subjects giving specified number of choices to picture 6GF.

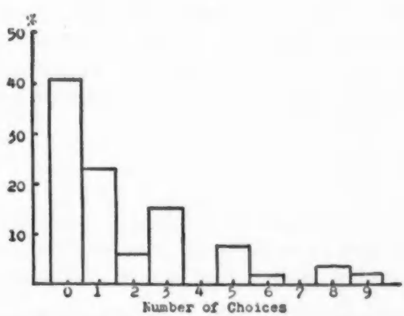

FIGURE 3. Percentage of subjects giving specified number of choices to picture 3GF.

as for picture 6GF, to a marked positive skew, in the case of picture 3GF. Evidently, then, agreement is closer in the case of pictures at either end of the happy-sad scale than in the case of pictures near the middle, which are more ambiguous with respect to the criterion of judgment.

One might ask whether any or all of the subjects were unable to discriminate between the pictures on the basis of the specified criterion and chose purely at random. A precise way of answering this question is to examine the number of circular triads in each subject's list of 45 choices.[6] If subjects chose purely at random their choices should contain a significant number of circular triads. With $N = 10$, the maximum possible number of circular triads is 40. The greatest number obtained was 13 ($P > 0.99$). This number is attributable to chance; hence we may conclude that no subject made choices on a purely random basis.

It may also be asked whether agreement among the subjects with respect to the relative happiness of the pictures was significant. For the group of 52 subjects the observed coefficient is $U = 1.128$. Testing the significance of this by the χ^2 method we have $\chi^2 = 593.04$ with 48 degrees of freedom, which is significant far beyond any ordinary significance point. For the group of 11 subjects, $U = 0.422$, giving $\chi^2 = 295.56$ with 61 degrees of freedom, which is also highly significant. We may conclude that the observed values of U could not have arisen by chance in a population in which choices were made at random.

In view of some significant sex differences reported below, the medians and semi-interquartile ranges of the distributions of choices of the men and of the women were computed separately. The median numbers of choices for each picture were in the same order with respect to magnitude as they were for the two groups combined, except that pictures 10 and 7GF were reversed by the men. Analysis of the choices suggests that this difference was due to inconsistency in choice. The medians for the two groups ranked in order of magnitude give a correlation coefficient of $\tau = 0.96$ ($P < .01$). From this we may conclude that there was no significant difference, and probably no real difference at all, between the men and the women with respect to the order of happiness of the 10 pictures.

The Stories

The average number of stories assigned to each point on the scale by the five raters is given in Table II. Examination shows that the distribution is negatively skewed and has its mode at $+1$. The average

[6]A circular triad is an inconsistency in choice of the kind where A is chosen over B, B over C, and C over A.

TABLE II

Distribution of Ratings of 520 Stories by 5 Raters

Rating	Total No. of stories placed at this interval	Mean
+3	118	23.6
+2	425	85.0
+1	820	164.0
0	495	99.0
−1	521	104.2
−2	164	32.8
−3	57	11.4
Total	2,600	520.0

correlation between the ratings of the different raters is $\tau = 0.57$ ($P<.01$), which indicates a significant degree of agreement.

Table III contains the medians and semi-interquartile ranges of the distributions of ratings of stories told to each picture. For each story, the value used was the average of the ratings assigned to it by the five raters.

The medians tend to decrease with the decreasing happiness of the pictures. The correlation between the pictures, ordered in terms of happiness, and the median ratings of the stories, ordered in terms of magnitude, is $\tau = 0.78$ ($P<.01$). We may conclude, therefore, that this tendency is significant.

TABLE III

Medians and Semi-interquartile Ranges of Distributions of Ratings of Stories Told to Each of 10 Pictures

Picture	Mdn.	Q
8GF	1.17	0.84
10	0.58	1.00
7GF	0.75	0.58
1	1.00	0.68
6GF	0.29	0.79
5	0.28	0.82
7BM	0.50	0.81
9GF	0.12	0.88
6BM	0.00	0.95
3GF	−0.17	1.18

In Table IV are given the medians and semi-interquartile ranges for the distributions of response times and of the number of words used in stories for each picture. It will be noted that the median response times and the median number of words per story both tend to increase as the happiness of the pictures decreases. The correlation coefficients are $\tau = 0.56$ ($P < .01$) and $\tau = 0.64$ ($P < .01.$), respectively.

TABLE IV

MEDIANS AND SEMI-INTERQUARTILE RANGES OF DISTRIBUTIONS OF RESPONSE TIMES AND WORDS PER STORY FOR EACH PICTURE

Picture	Response times*		Words per story**	
	Mdn.	Q	Mdn.	Q
8GF	16.37	11.74	82.30	25.76
10	25.39	12.52	98.93	31.09
7GF	21.50	12.47	100.50	35.84
1	16.31	8.81	131.64	47.28
6GF	28.50	17.12	123.50	43.62
5	27.50	13.00	129.21	39.41
7BM	21.50	14.35	129.94	48.75
9GF	35.34	16.92	133.49	44.34
6BM	27.50	16.40	139.50	37.91
3GF	34.69	16.95	125.00	43.50

*Numbers refer to units of one second.
**Numbers refer to units of one word.

TABLE V

MEDIANS AND SEMI-INTERQUARTILE RANGES OF DISTRIBUTIONS OF RATINGS OF STORIES TOLD BY 26 MEN AND 26 WOMEN

Picture	Men		Women	
	Mdn.	Q	Mdn.	Q
8GF	1.33	0.88	0.93	0.94
10	0.50	1.18	0.50	0.82
7GF	0.83	0.54	0.58	0.68
1	1.07	0.48	0.88	0.84
6GF	0.62	0.86	0.08	0.63
5	0.50	0.87	0.17	0.62
7BM	0.50	0.76	0.50	0.89
9GF	0.38	0.78	−0.21	0.90
6BM	0.26	0.88	−0.30	1.00
3GF	−0.05	0.75	−0.62	1.26

The medians and semi-interquartile ranges of the distributions of the ratings of the stories told to each picture by the men and by the women are given in Table V. For both men and women the medians tend to decrease as the happiness of the pictures decreases. The correlation coefficients are $\tau = 0.78$ ($P<.01$) and $\tau = 0.73$ ($P<.01$), respectively.

It will also be noted in Table V that, for every picture except 10 and 7BM (where they are equal), the medians of the distributions for the men are higher than the corresponding medians for the women. Thus, while the happiness of the stories told by both sexes tends significantly to decrease with the decreasing happiness of the pictures, the men tell happier stories than the women for eight of the pictures. The significance of this difference was tested by ranking the 20 medians and correlating with the sex of the subject. Assigning a rank of 1 to the smallest median and a rank of 20 to the largest, and assigning a rank of 1 to the women and a rank of 2 to the men, the computed correlation coefficient was $\tau = 0.24$ ($P<.01$). That is, the men tell significantly happier stories than the women, irrespective of the picture to which the stories are told.

The medians and semi-interquartile ranges of the distributions of response time and words per story for the men and the women are given in Table VI. A comparison of the median response times for the sexes picture by picture indicates that the response time of the men is, without exception, higher than that of the women. Similarly, for all pictures the

TABLE VI

MEDIANS AND SEMI-INTERQUARTILE RANGES OF THE DISTRIBUTIONS OF RESPONSE TIMES AND WORDS PER STORY FOR 26 MEN AND 26 WOMEN

Picture	Response time				Words per story			
	Men		Women		Men		Women	
	Mdn.	Q	Mdn.	Q	Mdn.	Q	Mdn.	Q
8GF	18.21	10.25	14.74	11.60	89.00	27.48	73.50	22.35
10	27.50	21.50	22.50	7.55	101.26	36.96	83.50	21.50
7GF	28.00	16.25	20.97	15.73	113.36	39.88	95.90	27.96
1	19.24	8.25	12.44	8.08	146.00	40.01	109.50	33.19
6GF	32.75	17.60	23.91	14.36	143.50	43.34	106.75	37.52
5	32.39	17.06	14.50	10.56	148.38	35.22	113.50	35.42
7BM	28.50	14.00	17.06	13.58	154.50	55.25	109.50	29.80
9GF	39.42	16.66	29.50	17.06	141.10	40.50	120.17	44.28
6BM	35.70	18.85	21.85	12.41	141.10	40.80	133.50	41.25
3GF	40.93	10.26	24.43	15.14	138.50	33.91	109.83	50.80

men used a higher median number of words per story than the women. For both sexes the median response times and the median number of words per story tend to increase as the happiness of the pictures decreases. For men and women, respectively, the coefficients for response times are $\tau = 0.73$ ($P<.01$) and $\tau = 0.33$ ($P = .11$). For words per story the coefficients are $\tau = 0.30$ ($P = .15$) and $\tau = 0.75$ ($P<.01$). That is, the response time for the men increased significantly as the happiness of the pictures decreased, but this was not true for the women. Conversely, the women used significantly more words in stories told to less happy pictures, while the men did not.

It was noted above that the women tended to have shorter response times than the men. The significance of this difference was tested by assigning a rank of 1 to the shortest median response time and a rank of 20 to the longest, and assigning a rank of 1 to the women and a rank of 2 to the men, and correlating. The coefficient thus computed was $\tau = 0.51$ ($P<.01$). That is, women had significantly shorter response times than men, irrespective of the picture. It was also noted that the women used fewer words per story than the men. Similarly computed, the coefficient in this case was $\tau = 0.32$ ($P<.01$), that is, the difference was significant.

Discussion

A general conclusion warranted by these results is that, in this situation, *the structure of the pictures* (i.e. the stimulus or structural variable) *had a significant effect on the stories told to them.* The set created in the subject by the instruction to "make up as happy a story as you can to each picture" (i.e. the organismic or functional variable) did not affect the feeling tone of the stories to such an extent that they were unrelated to the feeling tone of the pictures. To generalize, *it seems probable that the structure of TAT pictures plays a greater part in determining the nature of the stories told to them than appears generally to be supposed.* Stein (3) and other authors, by listing common themes, implicitly support this conclusion with respect to the content of the stories.

One plausible explanation of the fact that the response times and the number of words per story significantly increased as the happiness of the pictures decreased is to attribute it to the increasing inconsistency between the subject's "set" and the stimulus. It became progressively more difficult to tell happy stories as the pictures became less happy. Thus the subjects took more time to think of a story and used more words trying to "explain away" the unhappy features of the picture. An

alternate explanation might be that these increases were due to some form of fatigue, since the pictures were presented in the same order to all subjects with the less happy pictures near the end of the session. It would be necessary to repeat the experiment, presenting the pictures in a different random order for each subject, to decide between these two explanations.

The significant sex differences observed suggest that stimulus and organismic variables do not play the same relative roles in the perceptions of all individuals. Rather, the women seem to have responded more directly to the stimulus than the men, so that their stories were more decisively affected by it. The stories of both sexes varied significantly in happiness with the happiness of the pictures, but it will be recalled that the stories told by the men were significantly happier, in an absolute sense, than those told by the women. The response times for the women were also significantly shorter than for the men, as was the total number of words used per story.

These sex differences are consistent with some of the results obtained by Witkin and Asch (4). Their findings lead them to conclude that there are genuine differences between the sexes in the dependence of perceptual processes upon the stimulus field, the processes of women being more dependent than those of men.

Summary

Ten TAT pictures, which had been ranked by judges from happy to sad, were presented to each of 26 men and 26 women subjects, with the instruction to "make up as happy a story as you can to each picture." The stories were then rated by each of five raters on a 7-point happiness-sadness scale. The main results may be summarized as follows:

(*a*) For the group as a whole and for both sexes separately, the happiness of the stories told varied significantly with the happiness of the pictures. Men, however, told significantly happier stories, in an absolute sense, than women.

(*b*) For the group as a whole, both response time and number of words per story increased significantly as the happiness of the pictures decreased. For response time, the increase was significant for the men but not for the women. For number of words per story the increase was significant for the women but not for the men.

(*c*) The men had significantly longer response times than the women and used significantly more words per story, irrespective of the picture.

These results are interpreted as indicating: (1) that the structural

features of the TAT pictures significantly affect the feeling-tone of stories told to them; and (2) that the perceptual processes of women are more dependent on the stimulus field than are those of men.

REFERENCES

1. Frank, L. K. Projective methods for the study of personality. *J. Psychol.*, 1939, 8, 389–413.
2. Kendall, M. G. *Rank correlation methods.* London: Charles Griffin, 1948.
3. Stein, M. I. *The thematic apperception test: An introductory manual for its clinical use with adult males.* Cambridge, Mass.: Addison-Wesley, 1948.
4. Witkin, H. A., & Asch, S. E. Studies in space orientation: IV. Further experiments on perception of the upright with displaced visual fields. *J. exp. Psychol.*, 1948, 38, 762–782.

INFLUENCE OF EARLY EXPERIENCE ON MAZE-LEARNING WITH AND WITHOUT VISUAL CUES

RONALD H. FORGUS[1]
University of Pennsylvania

The first paper (2) dealing with the effect of early experience on adult cognitive abilities reported that whether early visual or motor experience will affect the relative abilities of adult rats in solving problems at maturity depends largely on the relationship between the kinds of early experience and the requirements of the problem. For example, it was found that a group of rats who had extensive visual experience with inanimate objects learned to discriminate visual forms faster than another group of rats, who had extensive visual experience with the same objects, but who were also permitted to explore these objects physically. At first sight it would seem that the added motor exploration of the objects acted as an impediment to the solution of purely visual problems.

It was suggested, however, that it was not so much that the second group were inferior to the first in visual discrimination but that they were also responding more frequently to other non-visual "irrelevant" stimuli. It was postulated that the longer time taken to learn the visual discrimination problem could be attributed to the time it took to abolish the wrong responses to the irrelevant cues. Such motor cues would presumably be more dominant in the "Visual-motor" than in the "Visual" Group; and this would account for the longer time taken by the Motor Group in visual discrimination learning.

But what would happen if the problem could not be solved by visual cues but required responses to internal motor (i.e. kinesthetic) cues for solution? One would then predict that the group who had the extensive early motor experience should be superior. This is exactly what was found in our second study (3). In this experiment it was found that the Visual Group was superior in solving an elevated multiple T-maze when many visual cues were present, but that the Visual-motor Group was superior when the visual cues were greatly reduced; and these large differences were highly significant.

This second study demonstrated that motor experience can be used to advantage when the solution of the problem requires the utilization

[1]The author wishes to express his thanks to Dr. W. Shaw and Mr. George Croswell for their aid in the design, and building of the apparatus; to Mr. S. Glick for his help in collecting the data; and to Dr. F. Irwin for his suggestions which materially improved the text.

Canad. J. Psychol., 1955, 9 (4)

of motor cues. But it still does not answer the question whether or not the Visual-motor Group was inferior to the Visual Group in visual discrimination per se. The present study was designed to answer this question, and also to verify the results of the previous study with respect to the superiority of the Visual-motor Group in maze-learning under conditions of poor visual cues.

An answer was sought to the following question: What would happen if two groups of rats, one reared in a complex visual and motor environment and the other reared in a complex visual but relatively simple motor environment, were to learn, partially, a complex closed-alley maze with specific visual cues present, and were then tested for complete learning with these visual cues reduced? The group reared in the complex visual and motor environment was called the "Visual-motor" Group (Group 1); the second was called the "Visual" Group (Group 2). We specifically tested the following hypothesis: *There would be no difference between the two groups in preliminary learning with the aid of specific visual cues, but the Visual-motor Group would be superior to the Visual Group in completing the learning when these specific visual aids were removed.*

Method

Rearing

Two groups of male hooded rats were reared in two different environments from the age of 25 days until they were 85 days old. There were 14 rats in each group. The rearing conditions were identical to those used in the previous study (3), the report of which contains photographs of the two living cages. The cage which housed Group 1 was 5 ft. long, 5 ft. wide, and 15 in. high. The walls were painted a flat black, and there were white objects located within 15 in. from the wall all around the sides of the cage. The white objects consisted of wooden structures such as alleys, blocks, tunnels, and elevated platforms. Group 1 was called the Visual-motor Group since they had complex visual as well as complex motor experiences with the inanimate objects in their cages.

The cage which housed Group 2 had the same dimensions as the cage which housed Group 1 and the white objects were in the same relative position. However, these animals were only permitted to live in the inner 900 square inches of their cages. This was accomplished by inserting a plastic cage, 30 in. long, 30 in. wide, and 15 in. high inside a large cage identical to that of Group 1. Thus these animals could see the inanimate objects but were never permitted to traverse them. For this reason Group 2 was called the Visual Group.

Both cages had very good lighting during the day and contained food and water in the same relative positions.

Two facts about the living cages bear emphasis. First, the two groups were reared under similar conditions, especially with respect to visual experience with inanimate objects. Second, the greatest difference in the rearing conditions was that the Visual-motor Group had more opportunity for motor experience in negotiating a complex physical environment.

Procedure

The rats lived in their respective cages (except when temporarily removed for cleaning purposes) until they were 85 days old, at which time they were transferred to smaller laboratory cages, 21 in. long, 18 in. wide, and 10 in. high. At this stage of the experiment the rats of the two groups were mixed so that they could have the experience of living together with members of the other group. We used four cages, each cage housing seven rats, four from one group and three from the other group. The animals were housed in these cages until the experiment was completed.

The rats were placed on a 23-hour food deprivation schedule for the next seven days and were handled extensively each day. They were fed purina dog chow and lettuce for one hour every day. At the end of this week the preliminary test was begun. The animals were run in a large, well-lit room.

The maze used was an 11-unit closed-alley T-maze. The units were painted a flat black. Each arm of the T measured 15 inches and the width was 3 inches. The maze had sides which were 4 inches high. There was ¼-inch wire mesh across the top of the maze. The starting box and the goal box had doors for putting the animals in and taking them out. We could block each bifurcation point after the rat made his choice, to prevent backtracking. The entire maze was located on the floor.

The sequence of correct turns was: R *L* L *R* R *L* L *R* L *R* L *R* R *L* R *L* R *L* L *R* R (the italicized letters refer to forced turns). On any one trial there was a maximum of 11 possible errors. At the entrances to the two alleys leading from each bifurcation point visual aids were provided in the form of light cards through which the animal could easily push. One card contained white *horizontal* striations on a black background, the other card contained white *vertical* striations on a black background. The horizontally striated card always led to the correct choice whereas the other card always led to the blind. If the animal chose the correct path he would of course be able to go on to the next "T." If he chose the incorrect path he would have to come back, push through the "wrong" door again and go through the other door.

The animals were given one trial per day after being deprived of food for about 23 hours. After they made the 11 correct choices they were allowed to nibble at some purina dog chow for 15 seconds in the goal box. At the end of the daily run the animals were fed purina dog chow and lettuce for about an hour. They were never deprived of water. Each rat received 10 trials under these conditions. By the tenth trial each rat reached the criterion of two or less errors on two consecutive trials. Following the suggestion of Ritchie *et al.* (7) we did not give the animal too many trials during the preliminary learning period. These psychologists have hypothesized that the initial period of maze learning proceeds primarily on the basis of visual cues and is later guided by execution of kinesthetic "chain" responses. Assuming that this hypothesis is correct we wanted to begin the second critical test before the kinesthetic "chain" responses were established, otherwise the removal of the visual cues would have produced little difference in the learning of the two groups.

The day after each animal had received 10 trials the second and critical test was begun. During this test the cards with the striations were removed, so that the animal would complete its learning without the aid of this visual cue. As before, each animal was given one trial per day. All other conditions were the same as during the 10 preliminary trials. This critical test ended for each animal when he reached the criterion of two consecutive errorless trials.

The units of the maze were randomly interchanged from day to day during both preliminary and critical tests to eliminate the use of olfactory cues by the rat. Since

there were no constant sounds localized in any fixed positions we may assume that auditory cues were also eliminated. If we accept Honzik's (4) conclusions that a rat learns almost nothing when visual, auditory, olfactory, and kinesthetic cues are all removed, we must conclude that during the first test the animals were learning primarily on the basis of visual and kinesthetic cues, whereas during the critical test they were *completing* learning primarily on the basis of kinesthetic cues, although visual and tactual cues were probably also used.

Results

Error and time scores were recorded for performance on both the preliminary and the critical tests. Figure 1 shows the mean number of errors made by each group during each trial of both tests. There was not much difference between the two groups during the 10 preliminary trials, although Group 1 made slightly fewer errors than Group 2. When the visual aids were removed from trial 11 on, the errors for both groups almost tripled. For the next 10 trials both groups reduced the errors at about the same rate. However, after trial 20 Group 1 went ahead and reached the criterion at least seven trials earlier, on the average, than Group 2 did.

The data were analysed separately for the preliminary test and the critical test. The relevant statistics for the number of trials to criterion

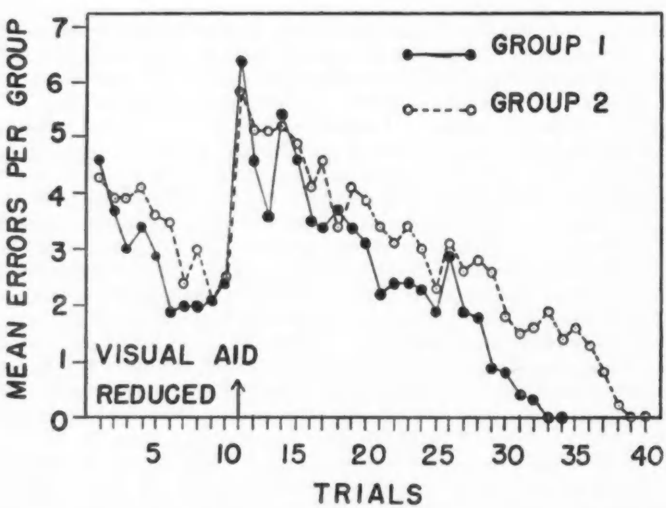

FIGURE 1. Graph showing the mean number of errors per trial for each group for the entire experiment, from the first preliminary trial to the last critical trial. Group 1 is the Visual-motor Group and Group 2 is the Visual Group.

and the number of errors made by each group on the preliminary test are shown in Table I.

From Table I we can see that the differences between the two groups in the mean number of errors and mean number of trials to the preliminary criterion were not significant.

It was also found that the mean times of the two groups were not significantly different. Festinger's non-parametric technique gave a *d* value of 1.14 ($p > .05$).

The relevant statistics of the analysis of the critical test scores are presented in Table II.

It can be seen that Group 1 took an average of about seven fewer trials to reach the criterion than Group 2. This difference is highly reliable ($p < .001$). The difference between the number of errors made by the two groups before the critical test criterion was reached is also significant, Group 1 making fewer errors than Group 2. Wilcoxon's unpaired replicates test of the difference gave a T of 15, significant at the 2% level.

TABLE I

Number of Trials and Number of Errors made by Each Group before reaching Preliminary Test Criterion

Statistic	Trials to criterion: Group 1 (Visual-motor)	Trials to criterion: Group 2 (Visual)	Errors: Group 1 (Visual-motor)	Errors: Group 2 (Visual)
Mean	6.43	7.36	27.79	33.36
Range	3–10	2–10	15–39	18–42
Standard deviation	2.09	2.34	6.61	7.70
t	1.06		1.98	
p	$> .05$		$> .10$	

TABLE II

Number of Trials and Number of Errors made by Each Group before reaching Critical Test Criterion

Statistic	Trials to criterion: Group 1	Trials to criterion: Group 2	Errors*: Group 1	Errors*: Group 2
Mean	19.50	26.64	60.50	83.14
Range	10–24	19–30	20–88	53–118
Standard deviation	4.049	3.904	—	—
t	4.58		—	
p	$< .001$		.02	

*As variances of the two groups were not homogeneous, Wilcoxon's T test was used to analyse the error differences.

As in the preliminary test, the time scores of the two groups were not different on the critical test. The Festinger *d* value was only .71 ($p > .05$).

It might be argued that differences between the animals on the critical test were affected by or related to their experience during the preliminary test. For example, it can be seen from Figure 1 that Group 1 made fewer errors during the preliminary test also, although this difference was not significant. To test this argument, a Kendall tau was computed to see whether there were significant rank correlations between animals' error scores on the preliminary and the critical tests, and between their trials to criterion scores on the preliminary and critical tests. The obtained taus of .23 and .12 were not significant ($p > .05$ in both cases). We may thus conclude that there was little relationship between performance on the two tests.

Discussion

First, we may conclude that the hypothesis tested was confirmed, that is, the two groups were not significantly different in the preliminary learning of the closed-alley maze with visual aids present, but the Visual-motor Group was superior to the Visual Group when learning was completed without the presence of these visual aids.

The first fact, namely, that the two groups were not significantly different in learning with the aid of visual cues, is consistent with the "perceptual learning" theory of Hebb (4), with the previous experimental findings of Hymovitch (6), and with those of the present author (2, 3) discussed in the introduction. Remembering that all tests in the previous studies were conducted in an open elevated maze with many "irrelevant" extra-visual stimuli present, we predicted that if the situation were made less complex, so that visual stimuli were the predominant cues, we would find no significant differences between the two groups. This is precisely what we found in the first test of the present study.

The fact that the complex early motor experience of an animal seems neither to improve nor to impair his visual discrimination per se fits in with the developmental theory of Hebb. Hebb takes the view that perceptual discrimination, excluding the discrimination of simple unity, is largely learned, and that the learning is governed primarily by perceptual experience.

But we must ask one important question. How general is this improvement in cognitive ability? Does it transfer to other cognitive processes? Hebb has suggested that it does. His position has been that the richness of perceptual experience determines the development of perceptual discrimination, problem-solving behaviour, and general intelligence. Very little has been said about the complexity of early motor

experience, however. The somewhat different findings of Hebb's students, Forgays (1) and Hymovitch (6), have not settled the question as to the importance of early motor experience as a determiner of adult problem-solving ability in rats.

Our own position has been that the relative importance of early visual and motor experience will depend on the nature of the problem requirements. To summarize, we previously found that animals with complex early visual-motor experience took longer than animals with primarily complex visual experience to solve primarily visual problems, when the stimulating field contained many irrelevant extra-visual stimuli as well as the relevant visual cues. Secondly, we found in the present study that this early motor experience is unimportant when the problem is best solved by utilizing visual cues, with irrelevant cues absent. And thirdly, we found in both studies that, when the visual cues are greatly reduced, the more complex motor experience did aid learning of the maze.

We may thus conclude that early visual experience improves discrimination and problem-solving ability in situations where visual cues are the primary aids to solution of the problem, and that complex motor experience improves learning and problem-solving ability in situations where kinesthetic cues seem to be important for solution of the problem. In other words, the relative influence of early experience on adult cognitive abilities depends largely on the relationship between the kind of early experience and the requirements of the problem task.

Summary

This study investigates the influence of early visual and motor experiences on the rat's ability to learn a complex closed-field multiple T-maze with and without visual aids. Two groups of male hooded rats were reared under different conditions from weaning until they were 85 days old. The Visual-motor Group was reared in a large complex environment which afforded much opportunity for visual and motor experience, whereas the environment of the Visual Group offered the same opportunity for visual experience but relatively little opportunity for motor experience with inanimate objects. As adults these rats were tested for closed-alley maze learning ability with and without "extra-maze" visual aids. The two groups were not significantly different in preliminary learning with the aid of visual form cues, but the Visual-motor Group was superior in completing learning when the visual aids were removed. The results are discussed in connection with previous

findings and with Hebb's developmental theory of the cognitive processes. It is concluded that the relative importance of early experience in determining adult problem-solving abilities depends on the relationship between the kind of early experience and the requirements of the problem task.

REFERENCES

1. Forgays, D. G., & Forgays, J. W. The nature of the effect of free-environmental experience in the rat. *J. comp. physiol. Psychol.*, 1952, **45**, 302–312.
2. Forgus, R. H. The effect of early perceptual learning on the behavioral organization of adult rats. *J. comp. physiol. Psychol.*, 1954, **47**, 331–336.
3. Forgus, R. H. Early visual and motor experience as determiners of complex maze learning ability under rich and reduced stimulation. *J. comp. physiol. Psychol.*, 1955, **48**, 215–220.
4. Hebb, D. O. *The organization of behavior*. New York: Wiley, 1949.
5. Honzik, C. H. The sensory basis of maze learning in rats. *Comp. Psychol. Monogr.*, 1936, **13**, No. 65.
6. Hymovitch, B. The effects of experimental variation on problem solving in the rat. *J. comp. physiol. Psychol.*, 1952, **45**, 313–326.
7. Ritchie, B. F., Aeschliman, B., & Pierce, P. Studies in spatial learning: VIII. Place performance and the acquisition of place dispositions. *J. comp. physiol. Psychol.*, 1950, **43**, 73–85.

ELECTROMYOGRAPHIC GRADIENTS AS INDICANTS OF MOTIVATION[1]

ALEXANDER K. BARTOSHUK
Allan Memorial Institute of Psychiatry and McGill University

SEVERAL recent papers (1, 10, 12, 13, 16) have reported on electromyographic (EMG) gradients: a progressive increase in muscle potentials from the beginning to the completion of a psychological task. Historically, these studies of EMG gradients were preceded by observations of gradients in grip and point pressure. Bills (2) in 1929 reported rising gradients in point pressure during arithmetic tasks, and suggested they were due to quantitative sets and compensating mechanisms for fatigue. McTeer (9) in 1933 noted gradients in grip pressure, especially if subjects were punished for errors (i.e. when motivation would have been highest). Telford and Swenson (15) in 1942 observed qualitatively that grip pressure tended to increase during mirror tracing, but they did not attempt a systematic explanation.

The earliest observation of a progressive rise in EMGs during mental tasks was apparently that of Davis (4) who called this the "build-up" phenomenon. However, variations in task difficulty precluded an adequate explanation of these forearm EMG gradients. Malmo *et al.* (10) in 1951 described qualitatively a tendency for neck and forearm EMGs to increase progressively during mirror tracing, and suggested that this might be an example of Hull's "goal gradient" phenomenon. As this observation was incidental to their experiment, their conclusions were perforce limited. It did, however, initiate a systematic study of EMG gradients. Smith (12) in 1953 presented quantitative evidence for EMG gradients in the active forearm extensors during mirror tracing. Bartoshuk (1) confirmed earlier findings of EMG gradients in the active arm, and ruled out the possibility that these EMG gradients were artifacts of movement speed by demonstrating that there were no corresponding drawing speed gradients. His observation of EMG gradients in the chin during mirror

[1]This investigation was supported in part by the Medical Research and Development Division, Office of The Surgeon General, Department of the U.S. Army, under Contract No. DA–49–007–MD–70, and in part by a research fellowship from the Canadian National Research Council. It was presented at the Annual Meeting of the Canadian Psychological Association, Montreal, June 1954. This report is based on a thesis submitted to McGill University in partial fulfilment of the requirements for the degree of Doctor of Philosophy. Grateful acknowledgment is made of the interest and support of Dr. Robert B. Malmo, and the technical assistance of Mr. R. E. Quilter.

CANAD. J. PSYCHOL., 1955, 9 (4)

tracing likewise indicated that movement was not essential for the occurrence of EMG gradients. This was also demonstrated conclusively under isometric conditions by the observation of EMG gradients in chin (13, 16) and forehead (16) during attentive listening.

The Smith and Bartoshuk studies (1, 12) provided the first quantitative evidence relating steepness of EMG gradients to a psychological variable. In accordance with Lewin's theory that "tension" tends to persist after incompleted tasks, Smith (12) found greater maintenance of EMGs after task interruption than after task completion. The EMG difference between completion and interruption correlated significantly with the steepness of EMG gradients for all tasks combined (1, 12). Because of this finding, and because the expected differences between completion and interruption were easier to demonstrate in highly motivated subjects (8), Bartoshuk (1) inferred that steepness of these EMG gradients indicated strength of motivation.

This conclusion was tested in the present study by means of the following two experimental hypotheses. Firstly, if we assume: (*a*) that steepness of EMG gradients is a direct function of strength of motivation to perform a given task, and (*b*) that performance is likewise a direct function of this motivation, then steepness of EMG gradients should be positively correlated with performance. Secondly, if we assume that strength of motivation can be maintained approximately constant from trial to trial (by the use of instructions and rest periods), then steepness of EMG gradients should likewise tend to remain constant throughout these trials.

Method

Experiment I

Subjects.[2] Sixteen airforcemen; median age 25, median years of schooling 11.

EMG recording. EMGs between bipolar surface electrodes were amplified and recorded on an Offner Type D, ink-writing electroencephalograph. Vacuum tube integrators using the condenser-charge principle were connected in parallel with the EEG recording pens (1, 3). Sponge electrodes, 1.5 inches apart, were used in each of the following bipolar placements: (*a*) right forearm extensor; (*b*) left forearm extensor; (*c*) right forearm flexor; (*d*) chin. The S was grounded by a metal electrode attached to the right wrist (3).

Drawing speed recorder. A 20-gm. cylinder of lucite was adapted to hold a piece of writing lead. An otoscope lamp was attached to the top of this lucite cylinder and connected to two 1½-volt batteries in series. The light from the lamp was transmitted through the lucite and focused at the writing point. Every time this pencil passed predetermined points in the task, a photoelectric cell provided distinct "on-off" signals on the EEG. This furnished data on drawing rate and accurately synchronized EMGs with percentage of task completed (1).

---

[2]The cooperation of the Royal Canadian Air Force personnel who served as subjects and of the officers who assisted with arrangements is sincerely appreciated.

Task material. The Ss traced circles, screened from direct vision but visible in a mirror. There were two concentric circles (radii 50 and 55 mm., respectively) and S was required to trace a circle within these two boundaries. Starting points varied as follows: bottom, top, left, right. Bottom-top and left-right axes were perpendicular to each other. Each tracing had one starting point, with an arrow to indicate drawing direction (always clockwise). This procedure ensured that, in a block of four trials, both task difficulty and amount of flexion and extension were equally distributed throughout the entire task.

Procedure. Subjects were told that this was a study of muscle tension during learning of a new task. Instructions emphasized that S try to remain always within the two circles, and that he try to complete each tracing as quickly as possible. For each mirror tracing, EMGs were recorded during a 10-second pre-period, performance, and a 15-second post-period. Performance trials were separated by 30 seconds while *E* changed the drawing sheet in the katograph.

This procedure was followed in trials 1–16. Then S was told, for the first time, that in subsequent trials he would receive a weak electric shock as soon as he completed each tracing.[3] He was also given a sample of the shock, which was used in trials 17–20. Previous instructions to trace quickly and accurately were repeated before each trial, with the addition that S was to be shocked as soon as he completed each tracing. After trial 20, S was told that he would not be shocked again, and the shocking electrodes were disconnected. Four additional trials, similar to the pre-shock trials, were given for control purposes.

Experiment II

The Ss were nine college students (four male; five female). Median age was 19. EMGs were recorded from: (*a*) right forearm extensor; (*b*) right forearm flexor; (*c*) left forearm extensor; (*d*) forehead (*frontalis* muscle). Instructions and procedure were identical with those of Experiment I, except for the omission of trials with expectation of shock, and inclusion of control tracings without mirror. Experiment II included two identical testing sessions of 44 trials each on separate days. Trials 1–4 and 41–44 in each session were control tracings without mirror; trials 5–40 were mirror tracings. Five-minute rest periods were inserted after trials 20 and 32, respectively.

Averaging of Data

Mirror-tracing trials were combined into blocks of four trials each because this equated amount of flexion and extension throughout the task. All results are averages for such blocks of four trials.

Results

The following phenomena of EMG gradients were observed repeatedly. At onset of a mirror tracing EMGs rose above the resting level and continued to rise progressively throughout the task. Within five

[3]Electric shock was used in only these four trials in an attempt to establish avoidance gradients (**11**). Because it has no bearing on the main findings reported, it will not be referred to again. The only effect of expectancy of shock was to increase (*P* .05) the average performance level of EMGs in left extensor (shock applied to left hand). In other words, the whole EMG gradient, of left extensor only, was shifted upwards without altering its slope.

seconds after completion of each tracing the EMGs decreased appreciably. This pattern was present not only in the right forearm, which was moving during mirror tracing, but also in muscles where no overt movement was observed, such as left arm and chin.

The percentage of Ss ($N=25$) with EMG gradients was 90 ($P<.001$)

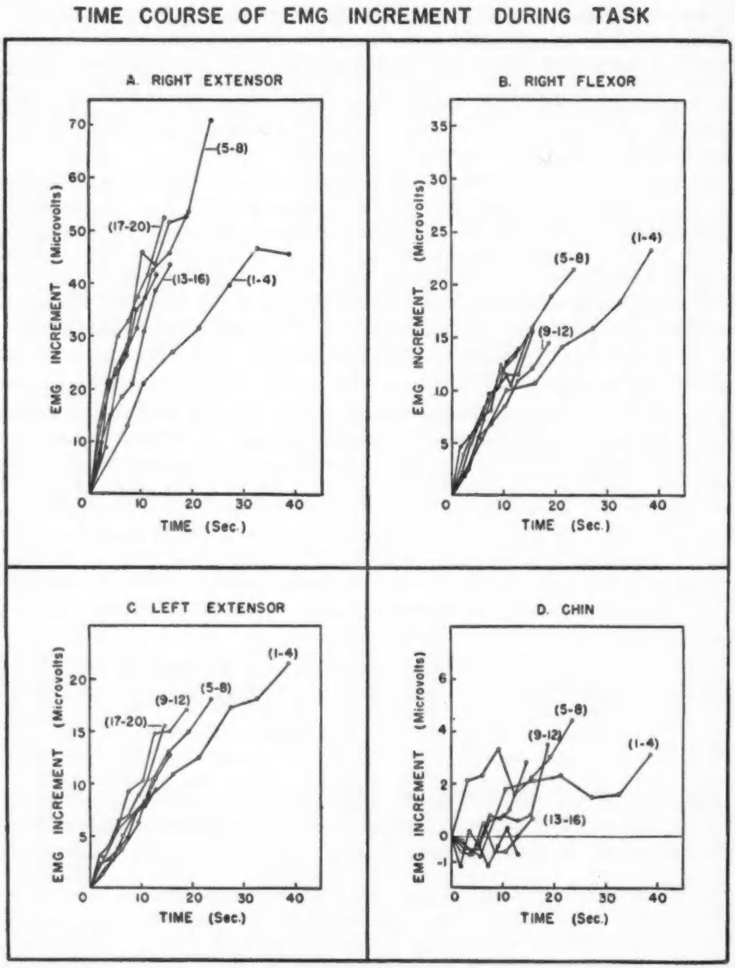

FIGURE 1. Vincent curves of mean rate of EMG rise. Each curve represents four trials; each point on the curve represents data from 16 Ss.

for right extensor, right flexor, and left extensor, throughout the 24 mirror tracings. Similarly, it was 90 per cent in eight Ss who performed 72 mirror tracings. For chin (16 Ss) EMG gradients occurred reliably in trials 5–8 (*P* .01) and 17–20 (*P* .05), and almost reliably in trials 1–4 and 9–12. Forehead EMGs recorded from nine Ss did not give reliable gradients.

Forearm gradients also occurred throughout 16 control trials of circle tracings without mirror.

Figure 1 answers the question: How fast in time did the EMGs rise? It was derived as follows. Each mirror tracing was divided into eight equal arcs. In order to omit large initial transients (in right arm) associated with change from rest to activity, the 12.5 per cent point of the circle was used as a reference point to calculate the subsequent amount of EMG rise during the task. This procedure gave, for each trial, seven pairs of measures consisting of an EMG increment together with its corresponding time interval. These data made it possible to assess the rate of increase in EMGs as a function of time. Figure 1 shows that the rise in forearm EMGs plotted against time gives an approximately straight line. Throughout all trials the EMG increments were almost linear functions of the drawing time increments. Therefore, slope of EMG gradients in any trial was expressed as average microvolt rise per second. This measure of slope was used throughout the subsequent analysis.

Analysis of Variance for Gradient Slope

An analysis of variance (Table I) revealed significant differences in slope of EMG gradients attributable to subjects (*P* .001), recording loci (*P* .001), and practice (*P* .05). The relation between slope of EMG gradients and recording loci was not affected by practice. Variations from person to person were seen in the form of the relationship between steepness of EMG gradients and recording loci (*P* .001), and also between steepness of EMG gradients and practice (*P* .05).

The relation between mean slope and practice was due to a significant (*P* .01) increase in gradient steepness from trials 1–4 to 5–8. Subsequently, throughout trials 5–24 there were no significant changes in group (N=25) means of forearm gradients. Similarly, group means did not change with practice for eight *Ss* who performed as many as 72 mirror tracings. These results agree at least partly with our initial working hypothesis that slope of EMG gradients should tend to be constant throughout the trials if we succeed in holding motivation constant. In apparent contrast, data available on 16 Ss revealed slight chin gradients which became less evident in later trials.

TABLE I

SUMMARY OF ANALYSIS OF VARIANCE IN SLOPE OF FOREARM EMG GRADIENTS

Source	Sum of squares	*df*	Variance estimate	*F* ratio
Loci (L)	654.20	2	327.10	8.72**
Practice (P)	26.58	5	5.32	2.30*
Subjects (S)	1,832.46	24	76.35	46.27**
L × P	29.81	10	2.98	1.81
L × S	1,799.38	48	37.49	22.72**
P × S	277.02	120	2.31	1.40*
L × P × S	395.48	240	1.65	
Total	5,014.93	449		

*Significant at the 5% level. **Significant at the 0.1% level.

The significant relation seen in Table I between slope and recording loci was due to the fact that EMG gradients in the right extensor were steeper (*P* .01) than in the right flexor or left extensor. There was no reliable difference between slope of right flexor and left extensor gradients. All forearm gradients were significantly steeper than the chin gradients (not included in Table I).

Consistent with these intermuscle differences was the finding that correlations between gradient slope in different muscles were generally low, except in the active arm (Table II). Slopes of gradients in right extensor were positively correlated (*P* .02) with flexor throughout all trials.

Intermuscle differences in gradient slope could not be readily attributed to variations in pre-level or "resting" level, because the cor-

TABLE II

CORRELATION BETWEEN EMG GRADIENTS IN VARIOUS MUSCLES

		Trials					
	N	1–4	5–8	9–12	13–16	17–20	21–24
Rt. extensor–Rt. flexor	25	.70**	.52*	.76**	.52*	.81**	.52*
Rt. extensor–Lf. extensor	25	.31	.40	.36	.37	.27	.39
Rt. extensor–chin	16	.17	−.04	.27	−.27	.10	.12
Rt. flexor–Lf. extensor	25	−.04	−.01	.33	.36	.20	.10
Rt. flexor–chin	16	.05	.23	.03	.06	.04	−.03
Lf. extensor–chin	16	−.09	−.07	.30	−.06	.29	−.12

*Significant at the 2% level. **Significant at the 0.1% level.

relation between slope of EMG gradients and pre-level also depended on the muscle studied (Table III). In right and left forearm extensors, slope of EMG gradients did not correlate significantly with pre-level. In right flexor, higher EMGs in the pre-period were generally followed by steeper EMG gradients. These positive correlations were not quite significant for trials 1–8, but significant (P .02) throughout trials 9–24. In the chin, higher pre-period EMGs were generally followed by less steep EMG gradients. These negative correlations were significant (P .05) throughout trials 9–20, and almost so (P .10) in trials 21–24.

Further analysis of the differences between subjects, which contributed markedly to the variance of gradient slope, revealed high intertrial correlations, that is, high individual consistency in slope of EMG gradients throughout 24 trials. Six blocks of four trials each yielded 15 intertrial correlations for each muscle. Median correlation (for 25 Ss) was: right extensor, .91; right flexor, .65; left extensor, .70; chin (only 16 Ss available), .54.

These correlations, together with the finding (Table I) that inter-muscle variations in gradient slope were related (P .001) to individual differences, indicate that subjects had a pattern of EMG response which was relatively consistent throughout the experiment.

TABLE III

Correlation of EMG Gradients with Pre-level

		Trials					
	N	1–4	5–8	9–12	13–16	17–20	21–24
Right extensor	25	−.25	.02	.16	.11	.08	.30
Right flexor	25	.38	.30	.55***	.51**	.51**	.60***
Left extensor	25	.21	−.11	−.16	.15	.27	.28
Chin	16	−.21	−.05	−.55*	−.66**	−.56*	−.46

*Significant at the 5% level. **Significant at the 2% level.
***Significant at the 1% level.

Relation of Gradient Slope to Performance

The major finding of this study is that the marked individual differences in gradient slope noted above were directly related to individual differences in quality of performance (when subjects were equated for amount of practice). Figure 2 shows the regression of performance scores (N=25) on slope of right extensor gradients for trials 1–4. Over a certain linear range, performance times and gradient slope were inversely

related; within this linear range the product moment correlation for 17 Ss was –.68. The other eight Ss had relatively short performance times and steep right extensor gradients in trials 1–4; but within this subgroup, performance times were not related to individual differences in slope of right extensor gradients. Since all performance times were at least 300 per cent of the minimum time physically required, drawing speed was not a limiting factor. Non-linearity in the relation of performance times to gradient slope was a genuine result of unusually steep gradients. Regression of error scores on slope of right extensor gradients was similar to that of performance times but less regular.

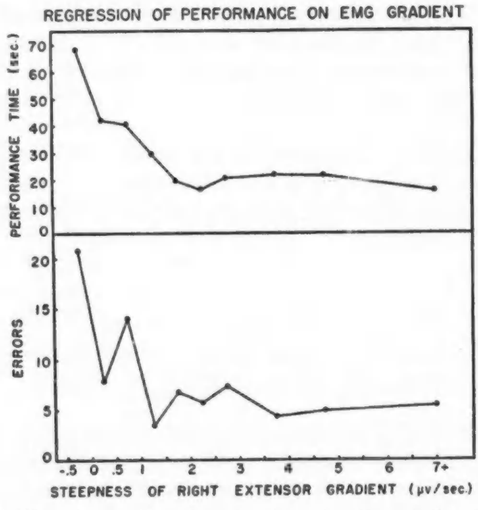

FIGURE 2. Non-linear regression of performance scores on right extensor gradient for 25 Ss in trials 1–4. Shortest performance times were at least 300 per cent of physical minimum.

Because performance scores were not a monotonic function of EMG gradients, and because of heteroscedasticity, the correlation statistics were not applicable. These results were analysed by separating the 25 Ss into two subgroups: Group I consisted of 12 Ss with greater than median steepness of EMG gradients in the particular trials being compared; Group II consisted of 12 Ss with below median steepness of EMG gradients. The results obtained from this analysis are summarized for each muscle.

Right extensor (*Table IV*). (*a*) Performance time was inversely related to gradient slope throughout 24 trials (P varied from .001 to .03).

TABLE IV

RELATION OF PERFORMANCE TO STEEPNESS OF RIGHT EXTENSOR GRADIENT

	Trials					
	1–4	5–8	9–12	13–16	17–20	21–24
Mean performance						
Time (sec.)						
Group I	19.37	15.93	13.58	12.01	11.31	10.80
Group II	47.38	24.08	20.22	16.93	16.31	13.74
p of difference	.001	.01	.03	.02	.03	.02
Performance time variance ($s2$)						
Group I	26.56	21.01	11.75	9.69	12.92	8.07
Group II	614.05	75.21	51.35	28.37	56.37	18.11
p of difference	.001	.10	.02	.10	.02	—
Mean No. of errors						
Group I	5.9	5.1	4.3	3.7	3.7	3.7
Group II	12.3	7.1	5.4	4.0	4.5	4.0
p of difference	.03	—	—	—	—	—
Error variance ($s2$)						
Group I	2.37	1.87	2.03	1.47	4.46	1.65
Group II	60.76	17.78	11.88	6.58	9.13	4.11
p of difference	.001	.002	.02	.02	—	—

Note. Group I consists of 12 subjects with greater than median slope, and Group II of 12 subjects with less than median slope of EMG gradient, in each block of trials.

(*b*) Variance of performance times was also inversely related to gradient slope (*P* varied from .001 to above .10). (*c*) Errors were inversely related to gradient slope (*P* .03) in trials 1-4; later trials gave only a slight trend in the same direction. (*d*) Error variance was inversely related to gradient slope throughout the initial 16 trials (*P* varied from .001 to .02).

Right flexor. (*a*) Performance time tended to be inversely related to gradient slope (*P* greater than .10). (*b*) Performance time variance was inversely related to gradient slope in trials 1–4 (*P* .002); similar unreliable trend present in later trials. (*c*) Errors were inversely related to gradient slope in trials 1–4 (*P* .03); similar unreliable trend present in later trials. (*d*) Error variance was inversely related to gradient slope in trials 5–8 (*P* .002) and trials 9–12 (*P* .02); similar trend occurred in other trials.

Left extensor. Performance time was inversely related to gradient slope in trials 1–4 (*P* .03) and trials 5–8 (*P* .01). Otherwise, performance scores were unrelated to gradients.

Chin (16 Ss). Gradient slope was inversely related (*P* .02) with error variance in trials 1–4, and inversely related (*P* .05) with number of

errors in trials 5–8 and 9–12. There were no other significant relations between chin gradients and performance.

Performance improvement rate. Individual differences in changes of gradient slope from trials 9–12 to 21–24 were directly related to individual differences in the corresponding intertrial improvement of performance times. Correlations were .42 (P .05) for right extensor and flexor respectively, and .35 (P .10) for left extensor.

These intertrial changes in EMG gradients and performance were both related to the third factor of performance in the initial four trials. Comparison of the 12 initially faster performers with the 12 initially slower Ss showed that the initially faster Ss maintained a slight performance edge throughout 24 trials. However, this initial advantage was reduced with practice. Between trials 9–12 and 21–24 the initially slower Ss showed greater absolute (P .01) and percentage (P .09) improvement in performance times than the initially faster Ss.

In the group of 12 initially faster Ss, average slope of EMG gradients increased with practice until trials 9–12, where these Ss had their steepest gradients. Subsequently, their gradients were positive, but less steep (Figure 3). Conversely, in the group of 12 initially slower Ss, average slope of EMG gradients increased with practice until trials 21–24. These group differences in changes of gradient slope from trials 9–12 to 21–24 were significant for right extensor (P .01), left extensor (P .02), and almost significant for right flexor (P .07). Consequently, from trials 9-12 to 21–24, EMG gradients increased in slope for the 12 initially slower Ss (whose performance improved most), but decreased in slope for the 12 initially faster Ss (whose performance improved least). This also illustrates nicely that greater absolute drawing speed can be associated with reduced slope of EMG gradients.

Role of physical factors. Several lines of evidence indicated that the above significant relations between gradient slope and performance scores were not artifacts of the following physical factors accompanying performance: (*a*) absolute drawing speed; (*b*) distance drawn; (*c*) fatigue due to physical work.

(*a*) Disparities between absolute drawing speed and slope of EMG gradients were observed in intratrial and intertrial comparisons. Within trials, drawing speed increased during the initial 12.5 per cent of each mirror tracing, remained approximately constant during the middle 75 per cent, and then decreased ($P < .001$) during the final 12.5 per cent of each tracing. This pattern occurred in all mirror tracings. EMG gradients clearly did not parallel these changes in absolute drawing speed. Moreover, absolute drawing speed increased 170 per cent ($P < .001$) from trials 5–8 to trials 21–24, whereas slope of right extensor gradients

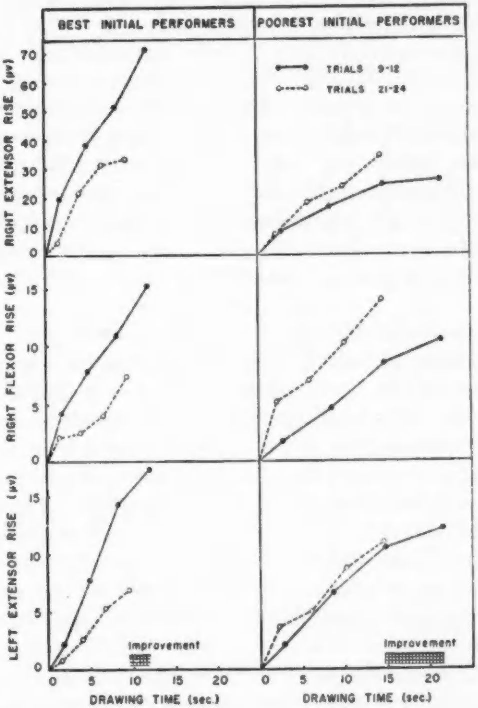

FIGURE 3. Relation of initial performance to correlation between rate of performance improvement and intertrial change in slope of EMG gradients. The 12 initially faster performers show only slight performance improvement from trials 9–12 to 21–24, accompanied by reduced gradient slope. Conversely, 12 initially slower Ss show marked performance improvement, accompanied by increased gradient slope.

decreased 10 per cent. These disparities between EMG gradients and absolute drawing speed (and others, e.g. EMG gradients in chin and inactive arm) indicate that EMG gradients were related primarily to factors other than absolute drawing speed.

(*b*) The marked, consistent individual differences in slope of EMG gradients could not be attributed to distance drawn, which was equal for all subjects.

(*c*) Fatigue (defined operationally either in terms of physical work done or impaired performance) was not a major factor, since the pencil was light (about 20 grams), the drawing periods were short, and there were many rest intervals. Moreover, positive correlations between slope

of EMG gradients and quality of performance rule out fatigue as a causal factor determing slope of EMG gradients.

Discussion

It is clear that the two main factors contributing to the variance in gradient slope were the differences between muscles of each subject, and the differences between subjects for each muscle. To account for the intermuscle differences would require a discussion of possible physiological mechanisms for generating EMG gradients, which will not be attempted at this point.[4] Instead, we will focus on the variations between subjects, and the hypothesis that they are indicative of differences in motivation.

In current research with humans there is no simple direct way to validate an indicant of motivational strength. However, by a hypothetico-deductive method, we can empirically test our postulated relations among: (*a*) the immediately antecedent environmental conditions, (*b*) the hypothetical motivational state of the organism (sometimes inferable from physiological measures), and (*c*) the observable behavioural consequences. It must be conceded that such functional relations are only aspects of motivation (i.e. partial definitions); but such quantitative relations between the motivational state of the organism (as inferred from physiological measures) and other empirical events are essential steps in the progressive delimitation of a theory of motivation by a series of approximations.

If, then, we assume that slope of EMG gradients is indicative of motivational strength, we should test our assumption by determining whether the slope is related to: (*a*) antecedent conditions expected to produce different intensities of motivation, and (*b*) the predicted consequences of varying strengths of motivation. This may be done in several ways. Available evidence has related gradient slope to (1) performance scores

[4]Clearly it would be desirable to have a physiological model which could explain the EMG gradients. Our data pose several problems: What mechanism could produce the slow progressive rise in EMGs which is of a much longer duration than the brief recruitment usually observed in muscle physiology? What determines the particular slope of EMG gradients, its variation from muscle to muscle, and apparent variation with S's motivational state? A physiological model must include not only the motor centres of the brain and their descending tracts which activate alpha efferents, but also the facilitating and inhibiting brain sites which modulate spindle activity (5) via the recently discovered gamma efferents (7). The gamma efferents–muscle spindles–monosynaptic arc–ventral horn cell form an external loop capable of modulating alpha efferents and maintaining muscle tonus. Hence this external loop, which is influenced by degree of muscle stretch and by suprasegmental factors, is one route whereby the central "arousal" system, referred to in Hebb's (6) theory of motivation, could facilitate responses, and possibly generate EMG gradients.

in the present experiment; and, in other experiments, to (2) greater maintenance of EMGs after task interruption than after task completion (analogous to the "Zeigarnik effect"); (3) magnitude of financial rewards offered S during tracking; and (4) reported interest in a recorded story or essay. The first two lines of evidence suggest how gradient slope, as an indicant of motivation, can support quantitative inferences about the consequences of motivation; (3) and (4) illustrate how antecedent conditions can be manipulated to induce different degrees of motivation, with corresponding changes in gradient slope. The argument for the data on interrupted tasks was presented in the introduction and elsewhere (1); we will now turn to our evidence on performance, and end with a summary of the remaining two studies.

Gradient slope was related to performance scores as predicted by our initial experimental hypothesis. Discussion of this hypothesis requires noting that in mirror tracing the improvement in performance speed and accuracy was due mainly to learning, rather than increased motivation. There are two points here. One is that, even if motivation were held constant, performance scores would improve owing to learning.[5] Therefore we must equate for practice when we try to relate performance scores to motivation. This was done by comparing Ss at the same stage of practice, assuming that more motivated Ss would have steeper EMG gradients and better performance scores than the less motivated Ss. Right forearm extensor gave the best results, satisfying this requirement entirely for performance speed, but only partly for accuracy scores. (This may be because the number of errors in later trials was low, making it difficult to assess the relation between errors and gradient slope.).

The second point is that, although performance would improve with practice when the group's average motivation was held constant, Ss whose motivation increased with practice should show more performance improvement than Ss whose motivation decreased with practice. This would be true whether motivation facilitated learning or merely performance. All three forearm placements (Figure 3) showed that an intertrial increase in slope of EMG gradients (indicating increased motivation) was related to greater performance improvement than that

[5]Consequently, by using performance scores at different stages of practice we had a way of rejecting the possibility that steepness of EMG gradients depended solely on the physical factor of movement speed, expressed as cm/sec. Because drawing speed increased markedly with practice, whereas group means for slope of EMG gradients remained approximately constant throughout most practice trials, we must conclude that the steepness of EMG gradients was not an artifact of the speed of movement. Moreover, the relation between EMG gradients and accuracy requires that we postulate some other factor, such as motivational strength, to account for the relations observed between performance scores and slope of EMG gradients.

associated with an intertrial decrease in slope of EMG gradients (indicating reduced motivation).

The fact that these later intertrial changes of EMG gradients were related to performance in the initial trials does not alter the argument; but it suggests that the Ss who were highly motivated initially became less motivated in later trials, whereas the Ss who were less motivated initially became more motivated as the session progressed.

The intertrial changes were generally not as marked as the differences between subjects which were maintained throughout all trials. Since these individual differences in gradient slope have to do with motivation, they may represent some enduring personality characteristic such as need of achievement.

It is also interesting to consider the non-linear relation seen in Figure 2 between performance and slope of EMG gradients (and therefore motivation). Within the limits of this experiment, it was demonstrated that, beyond a critical value, performance did not benefit from further increases in gradient slope. In fact, there is a suggestion that, for best performance, there is an optimal value of gradient slope, and that, for slopes in excess of this optimum, performance may be slightly impaired. Further study might provide data on the theoretical formulation that excessive motivation impairs performance (**6**).

In summary, the evidence that slope of EMG gradients is related to performance scores and to rate of performance improvement provides essential links between the hypothetical state of motivation (as inferred from physiological indicants) and its behavioural consequences.

This conclusion is supported by Surwillo's study (**14**) of EMG gradients during tracking under various degrees of motivation, induced experimentally by varying instructions and the magnitude of financial rewards S could earn for his tracking performance. Degree of induced motivation (as inferred from magnitude of the incentives offered) was statistically related to gradient slope.

Further confirmation is also provided by an experiment in which Ss were required to hear either a detective story or a philosophical essay played over three times (**16**). Of the four recording loci used, only the *frontalis* muscle gave consistently significant EMG gradients. In this situation, degree of induced motivation depended primarily on the story or essay stimulating the listener and arousing his interest. For the story, maximal interest was usually reported for the first hearing, decreasing with repetition; for the philosophical essay, maximal interest usually did not occur until the second hearing. For story and essay combined, *frontalis* EMG gradients were steeper in the most interesting hearing than in the least interesting hearing (P .03) regardless of order of pre-

sentation.[6] This agreement between steepness of EMG gradients and verbally reported motivation to listen complements nicely the other three pieces of evidence supporting the hypothesis that slope of EMG gradients is a direct function of motivation.

Summary

Previous electromyographic (EMG) studies of mirror tracing and attentive listening reported EMG gradients, that is, a progressive increase in muscle potentials from the beginning to the completion of a given psychological task. Results on completed and interrupted tasks suggested that the slope of EMG gradients may be indicative of motivation. If this were true, gradient slope should be directly related to performance speed and accuracy. This experimental hypothesis was tested in the present EMG study of 25 subjects through 24 trials of mirror tracing.

An analysis of variance revealed that significant differences in gradient slope could be attributed to intermuscle differences, and to differences between subjects for each muscle. With subjects equated for practice, gradient slope (especially for right forearm extensor) was directly related to speed and accuracy of performance. Moreover, intertrial changes in gradient slope were related to rate of performance improvement. Evidence from other studies reviewed showed that gradient slope was also related to reported interest in the task, and to magnitude of incentives offered for good performance. These results support the hypothesis that gradient slope is a direct function of strength of motivation to perform a given task.

REFERENCES

1. Bartoshuk, A. K. EMG gradients in goal directed activity. *Canad. J. Psychol.*, 1955, **9**, 21–28.
2. Bills, A. G., & Brown, C. The quantitative set. *J. exp. Psychol.*, 1929, **12**, 301–323.
3. Davis, J. F. *Manual of surface electromyography*. Montreal: Laboratory for Psychological Studies, Allan Memorial Institute of Psychiatry, 1952 (mimeo).
4. Davis, R. C. The relation of certain muscle action potentials to "mental work." *Ind. Univer. Publ. Sci. Ser.*, 1937, No. 5, 1–29.
5. Granit, R., & Kaada, B. R. Influence of stimulation of central nervous structures on muscle spindles in cat. *Acta Physiol. Scand.*, 1952, **27**, 130–160.
6. Hebb, D. O. Drives and the C.N.S. (Conceptual Nervous system). Presidential address, Division 3, APA, Sept. 6, 1954.
7. Leksell, L. The action potential and excitatory effects of the small ventral root fibres to skeletal muscle. *Acta Physiol. Scand.*, 1945, **10**, Suppl. 31.

---

[6]This finding is not reported by Wallerstein (**16**) and is based on my further analysis of his data.

8. Lewin, K. A. *A dynamic theory of personality.* (Trans. by D. K. Adams and K. E. Zener.) New York: McGraw-Hill, 1935.
9. McTeer, W. Changes in grip tension following electric shock in mirror tracing. *J. exp. Psychol.*, 1933, **16**, 735–742.
10. Malmo, R. B., Shagass, C., & Davis, J. F. Electromyographic studies of muscular tension in psychiatric patients under stress. *J. clin. exp. Psychopath.*, 1951, **12**, 45–66.
11. Miller, N. E. Comments on theoretical models. *J. Person.*, 1951, **20**, 82–100.
12. Smith, A. A. An electromyographic study of tension in interrupted and completed tasks. *J. exp. Psychol.*, 1953, **46**, 32–36.
13. Smith, A. A., Malmo, R. B., & Shagass, C. An electromyographic study of listening and talking. *Canad. J. Psychol.*, 1954, **8**, 219–227.
14. Surwillo, W. W. (Unpublished data, 1954.)
15. Telford, C. W., & Swenson, W. J. Changes in muscular tension during learning. *J. exp. Psychol.*, 1942, **30**, 236–246.
16. Wallerstein, H. An electromyographic study of attentive listening. *Canad. J. Psychol.*, 1954, **8**, 228–238.

PERFORMANCE ON SEVERAL CONTROL-DISPLAY ARRANGEMENTS AS A FUNCTION OF AGE[1]

MICHAEL HUMPHRIES AND ALFRED H. SHEPHARD
University of Toronto

THE EASE with which a complex motor task can be learned is frequently determined by the nature of the control-display relationship employed. Investigations of motor learning based on the modified Mashburn apparatus (**1, 2**), the Turret Pursuit (**5**), and the Two-Hand-Coordinator (**7, 8**) have demonstrated that adults have strong preference for certain control-display arrangements. Since there is no difference in the amount of physical energy required to operate a given control-unit for different control-display relationships, it has been suggested that the differences in ease of learning may be related to the effects of prior learning. Lewis and Shephard (**6**) hypothesized that habits acquired through everyday experience tend to facilitate performance on some control-display arrangements, and to interfere with others.

If, for a given display and control-unit, this hypothesis were tenable, and if any differential effects of prior learning were discounted, identical performance would be expected for all tasks. However, if everyday experience resulted in prior learning that tended to facilitate performance on one particular task more than on others, differences in performance, resulting from the facilitation and concomitant interference, should distinguish the tasks. Since the amounts of facilitation and interference from a habit are increasing functions of the strength of the habit (**2, 3, 9**), increases in the amount of prior learning should result in increases in the differences between the tasks.

If chronological age is taken as an approximate index of the amount of prior learning, the above deductions suggest that no differences in task difficulty, as measured by performance, should be noted in early life. As age increases, the effects of everyday habits should facilitate some tasks and interfere with others, resulting in an increasing trend in the differences in level of performance between tasks.

The present study was designed to investigate the relationship between chronological age and difficulty in performing three complex motor tasks. At each of three age levels, three groups of male subjects practised for seven minutes on one task of the Toronto Complex Coordinator.

[1]This work was conducted at the University of Toronto under Research Grant No. 265 from the Defence Research Board of Canada.

Method

Subjects

At each age level (5, 10, and 20 years), three groups were randomly assigned to the three tasks. This resulted in 9 groups of 10 subjects each, one group for each task at each age level. Male subjects only were used.

The 5-year-olds were obtained from day-care centres in the Toronto area. The 10-year-olds were local Cubs and Y.M.C.A. members. The adults were sophomores from the University of Toronto.

Apparatus

A complete description of the Toronto Complex Coordinator (TCC) has been given by Shephard (**10**) and Strahm (**11**). In many respects the TCC resembles the Modified Mashburn developed at the State University of Iowa (**4**). The subject is seated before a vertical panel of 81 light assemblies, each consisting of an outer red ring and a central green disc. Position of the illuminated red ring is determined by a 50-position stepping switch; that of the lighted green disc by S's movement of an airplane-type control stick. The task is to move the control to light the green disc within the illuminated red ring.

The type of electrical linkage enables *E* to vary the control-display relationship in either or both dimensions; for example, moving the stick to the right may move the green disc to right, left, up, or down, as desired.

Two measures of performance were assessed in the present study: number of "matches" and error-match ratios. A match was recorded each time the green light was correctly placed in the red ring. An error was recorded each time the green light was moved away from the red ring. The apparatus records errors for the vertical and horizontal directions separately, but for this report the two error scores were combined as a total error score. Error-match ratios were obtained by dividing total error score by total match score for each subject for each practice period. This provided an index of the average number of errors per match.

Measures were recorded in 20-second intervals, but were subsequently combined to give total values on each of the 7 one-minute practice periods.

Procedure

For any one age level the only difference between groups was in the control-display arrangement (task) employed. The three tasks were as follows:

Task A: Movement of the green light right, left, up, or down, resulted from stick movements right, left, away from, or towards the subject, respectively.

Task B: Movement of the green light right, left, up, or down, resulted from stick movements left, right, towards, or away from the subject, respectively. That is, the control elements had the same plane of movement as in task A, but the directional relationships were reversed.

Task C: In this task the planes were changed, so that movement of the control in the right-left plane moved the light in the vertical, whereas movement in the forward-backward plane moved the light in the horizontal plane. Movement of the green light right, left, up, or down, resulted from stick movements toward, away, left, or right, respectively. Preliminary studies had shown that this was the most difficult of the three tasks for adults.

Subjects were treated in the same manner for all tasks and age levels. They were told to make as many matches as possible, but at no time were they told that errors were being recorded. At the 5-year level it was necessary to read the

instructions more slowly, and in some cases to repeat them, but at no time were the 5-year-olds given additional information. All subjects practised without rest for a total of seven minutes.

Results

The data were analysed to determine the differences between trends for the three tasks as a function of age by: (*a*) an analysis of variance of each measure during the first minute of practice; and (*b*) a general description of the form of the curve for each group during the seven minutes of practice. The hypothesis, based on the transfer theory outlined earlier, suggested that performance on the three tasks would be approximately the same at the five-year level, and that, as age increased, the differences between tasks would increase for both measures of performance.

Figure 1 shows the mean number of matches for each group during the first minute of practice plotted against age. It is apparent that there was a difference in performance between the different tasks at the five-year level. Performance on all tasks improved with age, but the amount gained was not the same for all tasks.

Table I summarizes the analysis of variance for the data presented in Figure 1. There is a significant over-all trend ($F = 4.41$ with 4 and 81 *df*, $p < .01$) in the differences between tasks. From an inspection of Figure 1 it is apparent that the trend is an increasing function of age. However, partitioning of the degrees of freedom shows that the interaction was largely due to the characteristics of task C. There is no significant trend in the differences between tasks A and B taken alone. The trends in the differences between tasks A and C, and between B and C, are significant. This suggests that whereas all tasks improved with age, tasks A and B improved at approximately the same rate, whereas task C improved more slowly.

Figure 2 presents the mean of the error-match ratios for the first minute of practice on each task at the three ages. This shows a marked difference between tasks at the five-year level with a tendency for the tasks to become more alike as age increases. From Table II it can be

TABLE I

Analysis of Variance for Number of Matches

Source	Sums of squares	*df*	Mean square	*F*
Task	3362.466	2	1681.233	23.38**
Age	1620.466	2	810.233	11.27*
Interaction age × task	287.578	4	71.894	4.41**
Within groups	1319.500	81	16.290	

*Significant at the .05 level. **Significant at the .01 level.

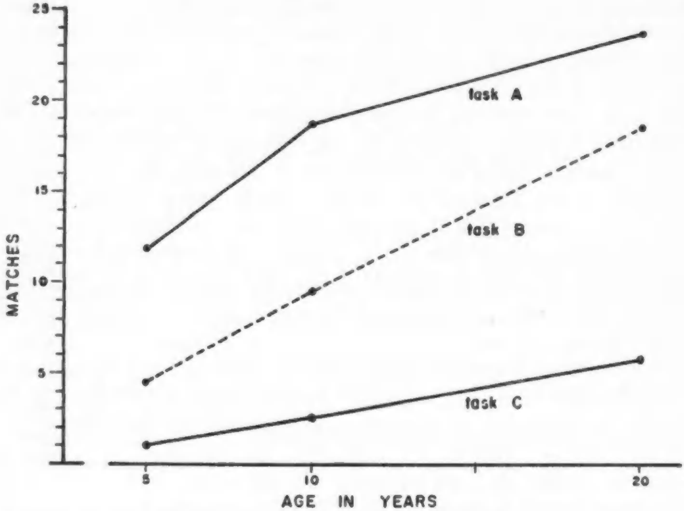

FIGURE 1. Mean number of matches for the first minute of practice as a function of age.

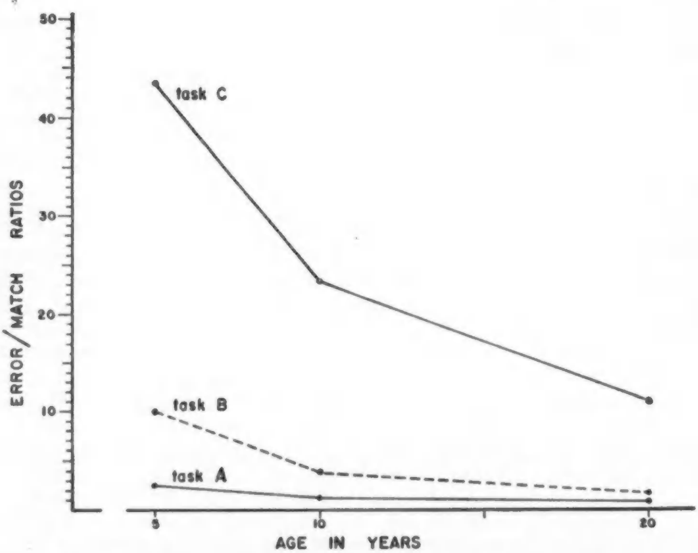

FIGURE 2. Mean error-match ratio for the first minute of practice as a function of age.

TABLE II

ANALYSIS OF VARIANCE FOR ERROR-MATCH RATIOS

Source	Sums of squares	*df*	Mean square	*F*
Task	10505.73	2	5252.86	7.762*
Age	3114.67	2	1557.33	2.301
Interaction age × task	2706.79	4	676.69	6.803**
Within groups	8057.12	81	99.47	

*Significant at the .05 level. **Significant at the .001 level.

seen that there is a significant trend in the differences between tasks ($F = 6.80$ with 4 and 81 *df*, $p < .001$). From Figure 2 it is apparent that this trend is a decreasing function of age, and is produced in the main by the characteristics of task C.

Figure 3 presents the mean number of matches made by each group for each of the seven one-minute practice trials. Although the data in this graph have not received statistical treatment, certain qualitative differences between the groups are apparent.

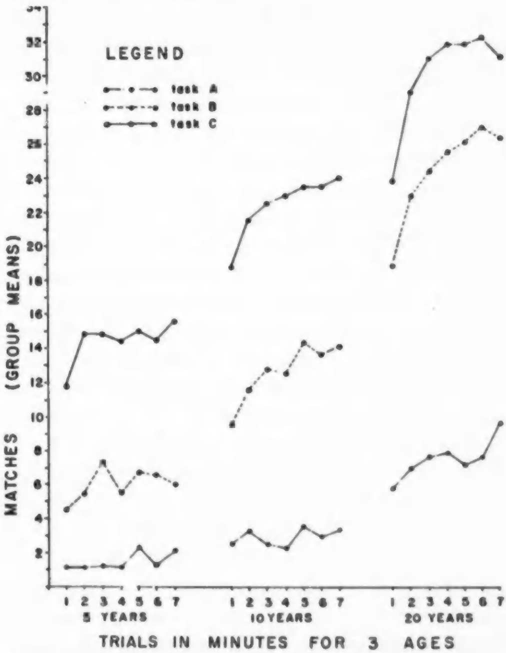

FIGURE 3. The relationship between age and performance on the three tasks.

In addition to the effects noted in Figures 1 and 2, it can be seen that performance on all tasks improves with age, and that at each age there is a difference between tasks. The slope of the curve for each task becomes steeper with age. In other words, the performance of older subjects improves more for the same amount of practice than does the performance of younger subjects. However, although not tested statistically, the slopes of the curves for tasks A and B appear to increase more rapidly with age than does the slope of task C. With successive age levels, task B shifts in position relative to the other tasks. At age five, it is closer in level and shape to task C than it is to task A. At the 10-year level it is midway between tasks A and C. At the 20-year level it is very similar to task A, both in shape and in absolute level. In other words, as age increases, task B shifts from resembling task C to resembling task A.

Discussion

The present study was based on a preliminary attempt to formulate and test certain hypotheses derived from the work of Lewis and Shephard (6). These suggested that adult preference for certain control-display arrangements of experimental apparatus is the consequence of habits acquired prior to the experiment, at least in those situations where there appears to be no physical reason for the observed differences in difficulty.

The present schema, based on transfer theory, states that differences in performance between tasks (control-display arrangements) having no inherent physical differences are attributable to perceptual-motor habits acquired through everyday experience. These facilitate performance on some tasks and interfere with (or facilitate less) performance on others. With increasing age, the amount of experience with the most usual control-display arrangements also increases; this increases the potential positive and negative transfer effects to new tasks. At maturity these effects result in preferences for certain control-display arrangements, which are not based on any inherent property of the tasks as such. As age increases there is a change in the relative difficulty of tasks which is attributable solely to associative or acquired factors.

Along with the changes in performance attributable to learning, there are certain changes related to variables subsumed under the general concept of maturation. As presently understood, the changes with age attributable to maturation facilitate all activities, at least up to the age of 30. Hence, if there is no innate preference for certain tasks, and if the effects of learning are discounted, maturation will have an equal and beneficial effect on performance on all control-display arrangements at a given age. As age increases, therefore, performance on all tasks should improve equally, but at each age there would be no difference in difficulty.

Where an initial difference between tasks can be demonstrated, and discounting the effects of prior learning, the gain in performance attributable to maturation should be the same for all tasks. Hence the absolute difference in performance between tasks should remain the same for all ages.

Where a trend in the differences between tasks is demonstrated, the differences should be attributable to learning factors. That is, if maturation affects performance on all control-display arrangements equally at any one age, any change in the difference between tasks as a function of age can reasonably be attributed to associative variables. This would be true whether an initial difference between tasks was demonstrated or not, and even though all tasks improved to some extent with age.

However, it may not be possible to assume that maturation affects all tasks to the same extent when they are originally different in difficulty. It would therefore be methodologically advantageous to demonstrate, at an age where performance was reliable, that the tasks selected were of equal difficulty. If it were then possible to show an increasing trend in the differences, even though all tasks improved with age, the transfer hypothesis would be tenable. Where it is impossible to find an age at which performance on the tasks is the same, it would be necessary to show, in addition to an increasing trend in the differences, a tendency for performance on some of the tasks to decrease with age (fewer matches and/or more errors in the case of the present study). If the transfer theory is supported by either of these methods it would then be possible to design studies to isolate the factors which determine the direction and nature of the transfer effects.

The data of the present study do not provide conclusive support for the above formulation of the transfer theory. For one thing, very large differences between tasks were obtained at the five-year level. Also, there was no indication that performance on any of the tasks decreased with age; in all cases, and by both measures, performance improved with age. Although there was an increasing trend in the differences between tasks, this might be explained by asserting that it reflects the differential effects of maturation on tasks which were, owing to some unknown factor, initially different in difficulty.

In spite of this, the transfer theory is probably more fruitful than that of maturation. The maturation hypothesis is, at present, too vague to be tested empirically. Furthermore, it remains to be shown that there is any inherent difference between the tasks other than those directly related to associative variables.

Summary

The present study was designed to investigate the relationship between chronological age and the relative difficulty of three control-display

arrangements of a complex motor learning apparatus. At each of three age levels, three groups of 10 male subjects practised one of three tasks on the Toronto Complex Coordinator. The three levels tested were at 5, 10, and 20 years.

At the 5-year level there was a significant difference between tasks. As age increased the differences between some tasks increased, while differences between others remained approximately the same. For all tasks the tendency to make errors decreased with age, although the rate at which they decreased was not the same. The results are discussed in terms of a proposed transfer theory.

REFERENCES

1. Adams, J. A. The evaluation of "difficulty" of task under several conditions of performance on the Modified Mashburn Apparatus. U.S. Navy, Office of Naval Research, *Technical Report* SDC 57–28, 1949.
2. Lewis, D., & McAllister, Dorothy, E. Facilitation and interference in performance on the Modified Mashburn Apparatus: II. The effects of varying the amount of interpolated learning. U.S. Navy, Office of Naval Research, *Technical Report* SDC 938–1–2, 1950.
3. Lewis, D., McAllister, Dorothy, E., & Adams, J. A. Facilitation and interference in performance on the Modified Mashburn Apparatus: I. The effects of varying the amount of original learning. U.S. Navy, Office of Naval Research, *Technical Report* SDC 938–1–1, 1950.
4 Lewis, D., & Shephard, A. H. Devices for studying associative interference in psychomotor performance: I. The modified Mashburn Apparatus. *J. Psychol.*, 1950, **29**, 35–46. (U.S. Navy *Technical Report* SDC 57–2–7.)
5. Lewis, D., & Shephard, A. H. Devices for studying associative interference in psychomotor performance: IV. The Turret Pursuit Apparatus. *J. Psychol.*, 1950, **29**, 173–182. (U.S. Navy *Technical Report* SDC 57–2–14.)
6. Lewis, D., & Shepard, A. H. Prior learning as a factor in shaping performance curves. *Proceedings of the National Academy of Sciences*, 1951, **37**, 124–131. (U.S. Navy *Technical Report* SDC 938–1–4.)
7. Norris, E. B., & Spragg, S. D. S. Performance on a following tracking task (modified SAM Two-Hand Co-ordination Test) as a function of the planes of operation of the controls. *J. Psychol.*, 1953, **35**, 107–117.
8. Norris, E. B. & Spragg, S. D. S. Performance on a following tracking task (modified SAM Two-Hand-Coordination Test) as a function of the relations between direction of rotation of controls and direction of movement of display. *J. Psychol.*, 1953, **35**, 119–129.
9. Shephard, A. H. Losses of skill in performing the standard Mashburn task arising from different levels of learning on the reversed task. U.S. Navy, Office of Naval Research, *Technical Report* SDC 938–1–9, 1950.
10. Shephard, A. H. The Toronto Complex Coordinator. *Canad. J. Psychol.*, 1955, **9**, 7–14.
11. Strahm, Carolyn L. The influence of instruction on performance of a complex perceptual motor task. *Canad. J. Psychol.*, 1955, **9**, 168–172.

ESTIMATION OF TRACK TARGETS AFTER PRE-VIEW

HUGH M. BOWEN AND MURIEL M. WOODHEAD
Medical Research Council, Applied Psychology Research Unit
Cambridge, England

PREDICTION of the final outcome of an incomplete two-dimensional track is made possible by familiarity with all the components of which the track is composed. This may be cited as an example of Bartlett's "preview" case of anticipation in human performance (1). In this type of anticipation, structural characteristics of a perceived pattern are retained to facilitate later perception or action or reproduction. In the present experiment the task was reproduction of a previously seen track.

The objective of this research was to evaluate four different displays in terms of the characteristic distortion which each imposed on a variety of tracks. The types of display simulated those used in certain actual navigational displays which give information in coded form rather than by direct representation. The four investigated were of the polar coordinate and cartesian coordinate type, each being considered with linear and exponential range projection. The test aim was to determine which display was more easily interpreted when the task was to predict the further course of tracks previously seen, but now shown with only the first third of each track visible.

On actual displays track information appears serially. Occasions arise in which estimation and decision have to be made as quickly as possible and the practised operator finds himself predicting the course of the track almost immediately after an initial alert. For example, in air interception procedure using a radar scope, it is essential to predict the future position of the target aircraft from a small amount of information known about its flight path. Since there is a choice of displays, it is possible that one could be proved more advantageous than another for the operator.

A pilot experiment preceded the present investigation. The observer was shown a card on which was drawn part of a track which pointed towards a target line. This part-track consisted of a number of dots (2, 4, or 8) and formed a portion of either a straight or curved track. The end dots of these part-tracks were placed at varying distances from the target line. The target line was numbered, so that the observer could indicate any position on it by stating a number.

The task consisted of the observer imagining the track extended to the target line and estimating (*a*) where the extended track would hit it and (*b*) the length of the imagined extension.

Results showed that the amount of initiating track (i.e. the number of

dots in the part-track) was not important for the two forms of estimation. The only effect that increasing the size of the part-track had was to reduce slightly the variance of the error in aim scores. The length of the imagined extension was typically underestimated, and this underestimation was correlated with the length to be estimated, so that the greater the length the more the underestimation. The interest in the present experiment was to transfer this type of task to actual navigational displays and to study the errors associated with each display.

Method

Materials

A master grid, which was a plain circle of 10″ diameter divided into quadrants at N, E, S, W, was drawn. On this, three different types of track, of 9-cm. length, were plotted. The first was a straight line, horizontal or vertical; the second was a right-angled line, turned at the 5/4–cm. division; the third was a straight diagonal line. They were drawn in four different arrowed directions of progress in each quadrant of the master grid.

There were thus 48 tracks which were transferred to each of the four displays, a total of 192 tracks. From these each subject did 48 test questions selected by Latin square designs to ensure systematic distribution of types of track, quadrants, and displays. Where position on a display is concerned, other researches in allied fields have shown there is rarely any significant difference due to this factor (5).

The four specialized grid displays used in the test were photostat black-on-white plane figures. The *linearly scaled polar coordinate display* was a circle, 5″ in radius. Between centre and periphery were outlined 4 range rings at 1″ intervals, the centre being zero. Radial (azimuth) lines appeared at 30° intervals. The *exponentially scaled polar coordinate display* was similarly constructed with the exception of the range rings. These were arranged so that the physical distance between range rings progressively decreased from the centre at the following distances: 1.82″, 3.09″, 3.97″, 4.58″, 5.0″.

In the *linearly scaled cartesian display*, a 5″ square was divided by a vertical centre line representing 0°. Equidistant lines were drawn to left and right of the centre line to indicate 30° intervals of azimuth, the vertical sides of the square being 180°. Mileage was shown by horizontal lines at ½″ intervals with the base as zero. The *exponentially scaled cartesian display* differed in the horizontal range lines which were at progressively decreasing distances, ascending from base, similar to the exponential polar coordinate display. Each of the four displays represented an area of 100 miles' radius.

Any track which could be plotted on the master grid could also be plotted on any of the displays. The same tracks took on different lengths and shapes in accordance with the different grids on which they were plotted. There was no notation on any of the displays.

Procedure

A full-length track as it would appear on one of the test displays was shown to the subjects for 5 seconds on a blank card. The card was then replaced by the appropriate test display on which was drawn the first one-third of the previously seen track. The subject imagined the complete track transposed on to the test

display in the place indicated by the start of the track, and he marked with a cross his estimate of its end.

This procedure simulates fairly closely actual practice. The showing of the whole track to an inexperienced observer beforehand was intended to make his task comparable to that of an experienced observer drawing on his memory store. In fact, however, showing the exact track simplifies the task, since even an experienced observer would be in some doubt as to the exact shape of a particular track compared with the shapes of the family of tracks immediately surrounding it.

Care was taken that the orientation of the track on the card and of the first one-third of the track on the test display were different, in order to prevent direct visual imprinting. Five seconds were allowed for answering and this appeared ample; nearly all answers were made in less time.

Scoring was done by measuring the deviation of the subject's estimate from the true termination of a track. This error score was calculated in three ways: actual distance in inches, distance translated into miles, and direction of error. Measurement by inches provided direct error scores, so that comparison between displays could be made, regardless of scale differences. Translated into mileage, errors took account of range scales; identical scores in inches on different displays were therefore not identical in mileage. Error by direction was categorized: above, below, left, right, of correct track ending. These directions were determined relative to the orientation of the track.

Subjects

Twelve inexperienced subjects were used—naval ratings with above average IQs on AH.4 Intelligence Test.[1] Subjects did 10 trial questions and it was ensured that the task was understood.

Results

1. The *mean error scores in inches* are shown in Figure 1. There were no significant differences between the two polar displays or between the

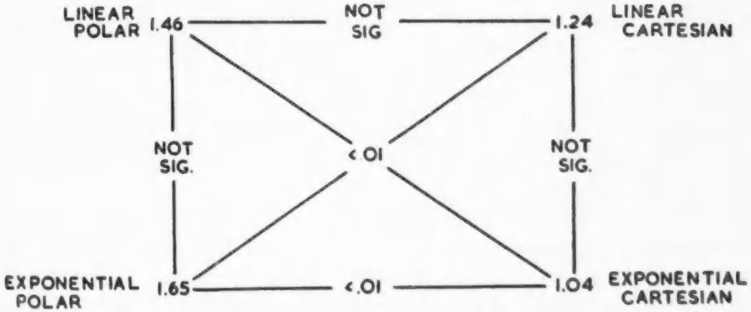

FIGURE 1. Mean error scores (inches) and P values of differences.

[1]The norms for this test were established on 3,407 adult factory population subjects. For further details see reference 4.

two cartesian displays. Of the other four comparisons possible, between the polar and cartesian grids, three gave significant differences ($p<.01$). These were

linear polar	from exponential cartesian
exponential polar	from exponential cartesian
exponential polar	from linear cartesian

The pilot experiments having indicated that length of track to be estimated affected accuracy of prediction, correlations were run between track lengths and error scores for the four displays (see Table I).

TABLE I

CORRELATION (PRODUCT-MOMENT) BETWEEN ERROR SCORE IN INCHES AND LENGTH OF TRACK FOR THE FOUR DISPLAYS

Scale	Coordinates	
	Polar	Cartesian
Exponential	+0.43 ($p<.01$)	+0.59 ($p<.001$)
Linear	+0.29 ($p<.05$)	+0.64 ($p<.001$)

All the correlations are significant, so it may be concluded that track length has a major influence on the prediction estimates.

However, since the same original tracks tended to be longer when plotted on the polar than when plotted on the cartesian displays, it might be thought that this effect explained the differences found between the polar and cartesian displays. To investigate this hypothesis an analysis of co-variance was made. The co-variance analysis for errors of track prediction, with track length kept constant, gave a significant result for the displays at $p<.025$ ($F = 3.34$ for dfs of 3 and 187). This indicated that something other than varying track lengths influenced performance on the four displays.

2. The *error score in miles* provided a measure of the actual type of error met in navigational practice. It should not be overlooked that the same physical size of error on the four displays could give very different mileage error scores. Figure 2 illustrates the mean error scores in miles. Results from the exponential polar grid were inferior to those from the other displays. It seemed unlikely that this was due solely to the exponential nature of the scale, because this factor did not result in differences between the two cartesian displays. Similarly, the polar coordinates

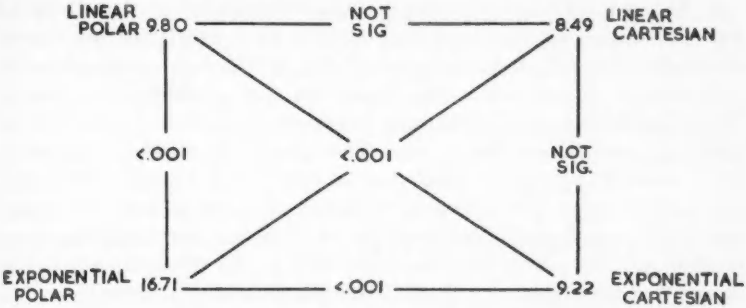

FIGURE 2. Mean error scores (miles) and P values of differences.

did not appear to be the sole cause, otherwise the linear polar display would have been expected to show significantly higher error scores than either of the cartesian. There was a strong supposition, therefore, of an interaction effect between type of display and type of scaling. Analysis of variance produced such an interaction term (F ratio sig. at .01 level), showing that the large error score was due to the combined effects of the polar coordinates and the exponential scale.

An examination of the standard deviations of scores on each of the four displays provided a further criterion for judging between them. Table II lists the standard deviations for each display, (A) in terms of miles and (B) in terms of inches. In the case of miles, the standard deviation for the exponential polar display was significantly greater than the other three standard deviations ($p<.001$ in all cases). In the case of inches, the standard deviations of the two exponentially scaled (polar and cartesian) displays were significantly greater than those for the two equally scaled grids ($p<.001$).

In operations where these displays are likely to be used, the range of judgments is frequently an important criterion. In general, large (even though infrequent) errors are unacceptable, and therefore it will be advantageous to use those displays which yield a small standard deviation in their range of scores. Both the exponentially scaled displays are inferior in this respect, the polar display particularly showing a large scatter when error is scored in miles.

TABLE II

STANDARD DEVIATIONS OF ERROR SCORES (N = 144 FOR EACH CONDITION)

	Linear polar	Exponential polar	Linear cartesian	Exponential cartesian
(A) Miles	5.67	10.31	6.11	6.54
(B) Inches	0.29	0.35	0.31	0.36

3. *Errors in direction* were categorized as shown in Figure 3. It can be seen that *underestimation* was the prevalent error, nearly 60% per cent of all observations falling in the quadrant labelled "below." A mean underestimation score was calculated, taking the sum of all the errors in the "below" quadrant, subtracting from it the sum of all errors in the "above" quadrant, and taking the average. This gave a figure of 9.7 miles for *mean underestimation.* This may be compared with a *mean error-of-aim* score of 10.5 miles (calculated by averaging all errors falling in the "left" and "right" quadrants). The mean underestimation score and the mean error-of-aim score were calculated for each of the displays, and it was shown that there was no difference between the displays for these two measures.

With errors measured in inches, a co-variance analysis for underestimations (calculated in the same way as for underestimations in miles), with track length kept constant, gave a non-significant result. This indicated that, after the effects of varying track length had been allowed for, there were no differences between displays for *underestimation* errors. This result confirms the finding that the amount of underestimation of a distance is a function of the length of that distance.

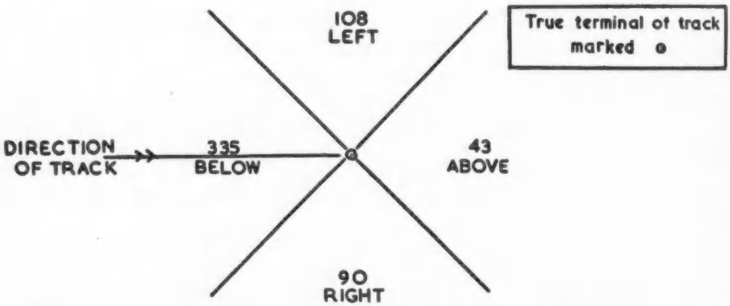

FIGURE 3. Errors in direction showing number of observations in each category.

DISCUSSION

It can be concluded that there are at least two factors subserving the prediction error. Firstly there is the underestimation error which is determined by the length of track to be predicted; secondly, there is a factor unconnected with length of track which operates differentially between the four types of display, and which is demonstrated by the error score in inches. This factor has a larger amount of influence on the polar displays than on the cartesian displays, as illustrated by Table I.

The error score in inches is more dependent on length of track in the cartesian than in the polar displays. We may speculate that there is a distortion factor for each display associated with the manner in which originally simple tracks are portrayed on the various display grids. What constitutes good or bad distortion for the human observer we cannot say from this experiment. Whatever it is, it seems to occur more on polar coordinate displays than on cartesian coordinate displays.

Results were carefully inspected for any learning effect, but in the three testing sessions which each subject had there was no evidence that his performance varied.

The original intention of the experiment—to investigate ease of interpreting four navigation-type displays when track prediction was the set task—is realized by inspection of the error score in miles. The exponentially scaled polar grid gives poorer results than the other three types of grid. The track distortion imposed by polar displays tends to lead to greater error than that resulting from cartesian displays. The size of the display will have a negligible effect upon the error because the error was found to be proportional to the length of track. That is, however long the track, the error will be approximately proportional to its length.

Fitts comments (3), in reference to maps, graphs, television screens, and many other projection surfaces, that few studies have been made of the ability of individuals to interpret rectangular and polar systems. Crook (2) also refers to the lack of research in problems of perception in relation to maps and charts. From the present experiment it seems that in navigation situations which require quick extrapolation from a present to a future position, the size of the display is immaterial. A cartesian or rectangular system seems to have certain advantages; and an exponential scale has disadvantages in that it gives a wider scatter of judgments. Certainly a polar display, exponentially scaled, should be avoided. However, it should be made clear that these results provide only partial criteria for judging the efficiency of a display, as other conditions and requirements have frequently to be taken into account.

Summary

An experiment was conducted to explore differences between four types of navigational display for a track prediction task.

Twelve subjects predicted from immediate experience the targets of initiated tracks on four grids: (1) exponential polar coordinate, (2) linear polar coordinate, (3) exponential cartesian coordinate, (4) linear cartesian coordinate.

Two forms of error were distinguished. One form was associated with the distance to be predicted and was typically an error of underestimation.

The other was associated with the type of display on which the prediction is made, and was greater with polar-type displays than with cartesian-type.

Of the four displays used, the exponentially scaled polar coordinate yielded significantly worse performance than the other three.

REFERENCES

1. Bartlett, F. C. Anticipation in human performance. In *Essays in psychology dedicated to David Katz*, pp. 1–17. Uppsala: Almqvist and Wiksells, 1951.
2. Crook, M. N. Printed materials, maps and charts. In *Human factors in undersea warfare*, pp. 61–76. Washington: National Research Council, 1949.
3. Fitts, P. M. Engineering psychology and equipment design. In S. S. Stevens (Ed.) *Handbook of experimental psychology*, pp. 1287–1340. New York: Wiley, 1951.
4. Heim, Alice W. *Manual for the group test of general intelligence, A.H. 4.* London: National Foundation for Educational Research (79 Wimpole Street, W.1.).
5. Williams, S. B. Visibility on radar scopes. In *Human factors in undersea warfare*, pp. 101–128. Washington: National Research Council, 1949.

RETINAL GRADIENT OF OUTLINE AS A STIMULUS FOR SLANT[1]

W. C. CLARK,[2] A. H. SMITH, AND AUSMA RABE[2]
Queen's University

It has been shown (1) that the retinal gradient of texture compression is a sufficient stimulus for the perceived slant of a surface in the absence of other cues for depth; also that, within certain empirically determined limits, there is psychophysical correspondence between the percept, the gradient of retinal stimulation, and the gradient of change in the structure of the physical surface (3). Since the retinal projections of the outlines of forms, textured or otherwise, generally follow the same rules of projective geometry as the retinal projections of units of surface texture (1; 3, pp. 82–83), retinal gradient of outline convergence should also be an adequate stimulus for slant when other cues are absent. Whether this is so does not appear to have been tested in this theoretical framework, although Gibson (3, p. 72) has suggested that linear perspective may be interpreted in terms of such changes in the proportions of figures in the visual field. The advantage of the approach is that it allows exact specification of the nature of the stimulus object.

The closest approach to an uncomplicated outline which we can produce experimentally is a "film-form," the area of retinal stimulation projected by a textureless physical surface with edges or a boundary. The bounded area has been termed a "film-colour" (6, pp. 7-12). Gibson's (4) "surfaceform" would seem to mean much the same thing, although it is not clear that he implies complete absence of visible texture. The edge or boundary of such a figure may be constituted by an abrupt change in brightness, hue, or saturation between the enclosed and enclosing areas (7, pp. 126–127). It has been argued convincingly that shape or form constituted by such a boundary is a geometrical abstraction which is perceived only at a given slant or orientation. This is, in part, the invariance hypothesis (7, pp. 222–223). The additional assumption involved is that the relationship is a regular one over a range of values of the variables, and on this point the evidence is inconclusive (8, pp. 875–877).

[1]This is the second of a series of papers on the role of monocular retinal cues in the perception of slant. The earlier studies were reported by Clark (1). This experiment was carried out under a grant from the Defence Research Board, Canada (Project DRB 370).

[2]Mr. Clark is now at the Vision Research Laboratory, University of Michigan. Miss Rabe is at the Department of Psychology, University of Michigan.

A further point of note about outline figures is that a given area of retinal stimulation which forms the basis for a judgment of shape may be produced by an indefinite number of stimulus forms, for example, the percept "rectangle" by different trapezia with the appropriate objective orientations. This has led to the belief, for which there is some evidence, that relatively simple retinal stimulation such as is provided by rectangles, circles, and the like, tends to elicit unstable percepts in the absence of the other visual and non-visual cues usually present in the pre-perceptual or learning stages. In a study of the nature of form (4) Gibson's ten observers all reported two or more, some as many as seven, different perceptions of simple outline drawings such as an ellipse, an arrow, and a trapezium, within a minute. Most of them reported seeing each form both in the frontal-parallel plane and in perspective. The explanation of these results was that "outline drawings are not appropriate stimulus objects with which to begin the study of form perception. They are habitually taken to *stand for* [italics ours] something other than what they are and, more important, what they stand for is often equivocal" (p. 410).

That impoverishment of stimulus situations should, per se, produce perceptual ambiguity, is at variance with the Gestalt point of view which requires that the shape-slant combination with the most stable organization, in Gestalt theoretical terms, be perceived. Ambiguous figures require special conditions of contour and figure-ground relations. Gibson's point in the study cited is different, namely that ambiguity arises from differences in the perceptual correlates of meaning, but there also seems to be an implication which is contrary to that writer's argument concerning the importance of texture gradients as stimuli for depth. It has been shown that both a regularly textured surface slanted from the frontal-parallel plane and a surface with texture density increasing from one edge to the other which is in the frontal-parallel plane elicit percepts of slant under conditions which yield little ambiguity. Gibson (3, p. 174) has suggested that the compression, rather than the slant, can be seen with special effort but that normally the slant is more compelling. If the theory has the generality which is claimed for it, it should also hold for outline drawings.

The aims of this experiment were to determine (*a*) whether or not monocular gradient of outline convergence is, like surface texture gradient, a sufficient stimulus for the perception of the slant of figured surfaces and (*b*) the relative stability or ambiguity of the related percepts.

Method

The apparatus consisted of the following: a head-holder and eyepiece permitting only monocular vision, hence eliminating cues of motion perspective and retinal

disparity; a reduction screen 32.8 cm. from the observer with a viewing aperture 11.7 cm. in diameter fitted with a shutter; a homogeneous black background placed in the frontal-parallel plane 175 cm. from O's eye; and stimulus forms, exposed by means of a holder 12 cm. from the screen, centred in O's visual field and thus surrounded by a black ground. Stimuli beyond 40 cm. from O are unlikely to elicit differential accommodation, especially if they have no microstructure (3, p. 112; 5).

The stimulus forms consisted of two rectangles, 28 cm. by 15.4 cm. and 19 cm. by 15.4 cm., and a trapezium with parallel sides 15.4 cm. and 14.2 cm., 25.4 cm. apart. The long axes of the film-forms were horizontal and they were slanted by rotation about a vertical axis with the left edge closer to O. The shape of the unslanted trapezium, that is the convergence of the retinal projections of its top and bottom edges, was approximately the same as that of the longer rectangle at 20° slant, and its shape at 20° slant was approximately the same as that of the longer rectangle at 40° slant. Homogeneous retinal stimulation was produced by using surfaces with as few and slight irregularities as possible (painted flat white), then decreasing the illumination well below the point of visibility of any remaining irregularities. The reflectance from the surfaces of the room was so arranged that there were no appreciable gradients of light and shadow on the stimulus surfaces.

There were ten stimulus conditions: each of the three film-forms exposed at physical slants of 0°, 20°, and 40° with the left edges closer to O, and a "reversed trapezium," produced by slanting the trapezoidal film-form 40° from the frontal-parallel plane in the opposite direction so that the shorter of the parallel sides was closer to the eye, a condition which produced a rectangular retinal image.

The comparison variable was a metal rod pivoted about a vertical axis on a circular dial placed on the table 20 cm. below the viewing aperture so that O could see it with very little head movement and adjust the pointer by means of two cords to a position which he judged to be parallel with the stimulus surface. The observer's reference dial showed only a cross, marking the four quadrants, but the device was coupled to a concealed pointer and dial graduated in degrees on E's side of the reduction screen.

Each of 7 college students made four 3-second observations with his dominant eye for each of the 10 stimulus conditions, in a random order of stimuli which differed for each O. Judgments were recorded to the nearest degree. The stimuli were concealed from O by the shutter while he was adjusting the comparison variable between trials. The observer returned the pointer to the frontal-parallel position after each trial. At the end of a session each O was again shown the entire series of stimuli, and was asked to draw on separate sheets of paper the shape of each figure at each stimulus slant, and to name the form he had drawn.

RESULTS

The slanted rectangular film-forms and the trapezium in the frontal-parallel plane were perceived as slanted (Table I). Mean perceived slant increased generally with increase in physical slant. The unslanted rectangular forms were perceived as unslanted but the unslanted trapezium was perceived as slanted. The "reversed trapezium" when slanted at 40° appeared as unslanted and rectangular in shape. The mean perceived slants are substantially less than the corresponding physical slants, a common finding in these experiments, and probably due to the

TABLE I

MEAN PERCEIVED SLANT IN DEGREES ($N = 7$)

Physical Slant (degrees)	Film-form			
	Long rect.	Short rect.	Trap.	Rev. trap.
0	.14	−.57	7.89	—
20	6.96	4.89	16.07	—
40	16.96	17.32	28.43	.64

Note. Each figure in the body of the table represents the mean of 28 observations.

absence of other cues for depth. None of the mean differences in perceived slant for the two rectangles at the three physical slants was statistically significant (Table II). All the mean differences between the perceived slant for the trapezium and the rectangles were significant except that involving the short rectangle at 40°.

Other relevant comparisons not shown in Table II are as follows: the mean perceived slant of the "reversed trapezium" at 40° slant was not significantly different from that of either the long or short rectangle at 0° slant ($t = 1.19$, $P > .20$; $t = 1.81$, $P > .1$, respectively); the mean perceived slant of the trapezium at 0° slant was not significantly different from that of the long rectangle at 20° slant ($t = .75$, $P > .4$), but was significantly different from that of the short rectangle at 20° slant ($t =$

TABLE II

SIGNIFICANCE OF MEAN DIFFERENCES IN PERCEIVED SLANT ($df = 6$)*

Comparison	Physical slant 0°		20°		40°	
	t	P	t	P	t	P
Long and short rectangles	1.972	>.05	1.035	>.30	.153	>.80
Long rectangle and trapezium	6.164	.005	4.555	.005	3.745	.01
Short rectangle and trapezium	10.317	.005	3.572	.025	2.389	.10
Reversed trapezium and long rectangle at 0°					1.190	>.30
Reversed trapezium and short rectangle at 0°					1.810	>.10

*The t ratios were computed by the method of paired replicates (2, p. 105).

2.86, $P < .05$); the mean perceived slant of the trapezium at 20° slant was not significantly different from that of either the long or short rectangle at 40° slant ($t = .32$, $P > .7$; $t = .29$, $P > .7$, respectively).

The verbal reports obtained when the perceived shapes were drawn, rather than the measurements of the drawings themselves, were used with the above data as the basis for inferring the shapes which were seen. This was because errors of but a few millimetres in the judgments represented by the drawings produced very gross differences in the physical slants which were calculated to be equivalent to the drawn shapes.

Discussion

These results show that the retinal gradient of outline convergence is, in the absence of other visual cues, a sufficient stimulus for perceived slant. They therefore confirm the extension of Gibson's theory of psychophysical correspondence to gradients other than surface texture, although it is clear that gradient of outline is a closely related variable. Regardless of the orientation of the distal stimulus form (cf. the reversed trapezium at 40° slant), without convergence the retinal image inferred from the percepts was rectangular. Conversely, when the opposite edges of the retinal image converged, slant was perceived and increased as the retinal gradient of convergence increased, also regardless of the objective orientation of the stimulus form. This is indicated by the insignificant difference between the mean percepts for the trapezium at 0° and 20° slant and those of the long rectangle at 20° and 40°, respectively. If the correspondence is described as psychophysiological rather than psychophysical the nature of the relationship becomes clearer. It is perhaps more appropriate to describe the correspondence as being between inferred retinal gradient of convergence and perceived slant rather than between actual and perceived slant, since other cues for depth were excluded. The experimental conditions established that convergence in the retinal image was the critical variable.

The stimuli used in this experiment were film-forms rather than simple outline drawings such as Gibson's (4) subjects reproduced, but they were very similar to some of these drawings. However, no observer reported that the stimulus forms changed in their shape-slant relationships. They appeared quite stable, whether slanted or in the frontal-parallel plane. This is at variance with the view, required by a "meaning" theory of visual perception, that outlines of surfaces result in ambiguous percepts in the absence of other cues. The explanation may be that the stimulus forms in the present experiment were set off from the background so that the changes in their slant presented definite out-

line convergence cues, whereas Gibson's figures were drawn on grounds oriented in the frontal-parallel plane. This would be expected to produce in the observer a conflict between a slant percept correlated with an outline drawing and the "unslanted" percept required by the orientation of the ground.

Different subjective sets may also have been operating in the two situations, although this is less likely. Gibson's observers were simply instructed to tell what they saw on the card and to describe any changes therein. The relevant part of the instructions in the present experiment was as follows: "Keep a flexible attitude and do not expect any particular slant or shape. I may present you with nothing but flat, unslanted figures, or nothing but slanted figures, or some may be flat and some slanted; likewise, some figures may be rectangular and some trapezoidal in shape, or they may be all rectangular or all trapezoidal." Aside from the difference between the two situations and the particular explanation which is favoured, the present results do not support the conclusion that outline drawings are generally ambiguous. Reversible figures represent a special case in which the figure-ground relations are such as to set up hypothetical dynamic stresses in the psychophysiological systems which mediate the percepts.

Koffka (7, p. 124) reported an experiment by Metzger under conditions similar to those of the present experiment, with the exception of one variable (the presence of microstructure in the ground), which may account for Gibson's results as an alternative to a "meaning" theory. A small, unilluminated square (homogeneous film-form) was produced on an otherwise slightly textured or microstructured white surface under low intensity illumination. When viewed at an angle the patch presented a black trapezium to the retina and the observers reported seeing a trapezium, along with the neighbouring region, in the frontal-parallel plane. The microstructure of the enclosing surface apparently constituted the entire percept. In the present experiment, in which both the film-form and the ground were homogeneous, the trapezoidal film-form was perceived as a slanted rectangle. Both the ground and the enclosed figures in Gibson's stimulus situations were microstructured, but the apparent effect of the ground, predictable from Gestalt theory, was preeminent. Of course, the outline convergence cue in the present experiment was more marked than in Metzger's experiment, and this, in conjunction with enclosing contour, would also be expected to favour "good figure." These outcomes raise a question concerning the effect of convergence on perceived slant in isolation. One would expect that the figure would lose some of its film-colour, tend to assimilate with the ground, and appear less slanted.

The experiment was not intended to be a test of the invariance hypothesis, but there are some implications which merit comment. The following outcomes appear to accord with the invariance hypothesis: the trapezium at 0° slant appearing on the average slanted about the same amount as the long rectangle at 20°; the trapezium at 20° slant appearing as slanted about the same amount as the long rectangle at 40°; and the reversed trapezium at 40° appearing rectangular and frontal-parallel. However, some of the divergent individual results suggest the need for a detailed experimental definition of the limits of applicability of the hypothesis. The verbal reports, together with the foregoing data, indicate that individual percepts were generally stable. Almost all the observers saw slanted rectangles and slanted trapezia as slanted rectangles and slanted trapezia. However, the fact that some saw the unslanted trapezium as a rectangle and others saw it as a trapezium is unexplained by the present data and is not in accordance with the concept of an invariant relation between slant and shape.

Summary

Seven observers recorded their perceptions of the slant of rectangular and trapezoidal film-forms produced under low illumination by white figures against a homogeneous black ground. Essential conditions were monocular vision and fixed head at a distance which minimized or eliminated accommodation. The stimuli were presented in the frontal-parallel plane and at angles of 20° and 40° with the background. In these conditions retinal gradient of outline convergence was found to be a sufficient stimulus for slant, in accordance with the theory of psychophysical correspondence. With some individual exceptions, the results agreed with the invariance hypothesis. Shape did not yield ambiguous percepts.

REFERENCES

1. Clark, W. C. Monocular perception of slanted surfaces. Unpublished master's thesis, Queen's University, 1953.
2. Dixon, W. J., & Massey, F. J. *Introduction to statistical analysis.* New York: McGraw-Hill, 1951.
3. Gibson, J. J. *The perception of the visual world.* Boston: Houghton Mifflin, 1950.
4. Gibson, J. J. What is form? *Psychol. Rev.*, 1951, **58**, 403–412.
5. Ittleson, W. H., & Ames, A. Accommodation, convergence, and their relation to apparent distance. *J. Psychol.*, 1950, **30**, 43–62.
6. Katz, D. *The world of colour.* London: Kegan Paul, 1935.
7. Koffka, K. *Principles of Gestalt psychology.* London: Routledge and Kegan-Paul, 1935.
8. Stevens, S. S. *Handbook of experimental psychology.* New York: Wiley, 1951.

The Human Person: An approach to an integral theory of personality. By MAGDA B. ARNOLD AND JOHN A. GASSON. New York: Ronald Press, 1954. Pp. x, 593. $5.75.

TO DISCOVER a psychology text which includes a chapter on learning theory and also one on the spiritual exercises of St. Ignatius is a shock to the non-Catholic reader. It is to be hoped that such shock treatment will not result in a withdrawal response, but rather arouse interest in the relationship which the authors perceive between these apparently disparate phenomena.

In the late nineteen-thirties Magda Arnold was one of a group of graduate students who had tea in the cellar of the psychology department, University of Toronto. In those days there was violent rebellion against the artificiality of the ivory lab and the unreality of mechanistic personality theory. There was even a little low whispering that the clinicians were doing exciting things, and occasionally a surreptitious suggestion that Freud might possibly have had something worth considering. The search was for a psychology that described life in terms approximating the way in which human beings lived it, felt it, and perceived it; hence the group swung to an acceptance of some form of phenomenological approach: Psychological reality is that which is real to the person.

The essence of those basement tea parties has permeated into the psychological field through many channels, but Dr. Arnold has been unique in her progress since student days. She has been able, at least in her own mind, to expand this human point of view to embrace both the basic aspects of man's activity (particularly the physiology of his emotions, on which her studies are widely cited) and also the loftiest—his religious experiences and his magnificent attempts to solve the perplexities of abstract thinking in the realm of the spiritual. In Dr. Arnold's thinking there is no need for reconciling the physiological and the spiritual, for there is no conflict between them, but rather complete harmony. In this harmony it is made plain that the greater cannot be derived from the lesser; the lesser, however, must be included in the greater.

This book grew out of a workshop in Personality held at Barat College in 1951. Like most products of symposia it is uneven; while the chapters by Dr. Arnold and Father Gasson are always comprehensible, others introduce concepts and language foreign to the non-Catholic reader. It is stated that this book is written as a primary text for courses in personality.

It is doubtful whether it will serve such a purpose in non-Catholic colleges, but this reviewer would recommend it as required supplementary reading. It would be good for the modern student to learn that at least one group of psychologists is making a serious effort to include man's highest aspirations and achievements in the basic data of our science, and to account for these human characteristics without considering them derivatives from, or compensations for, some animal need.

Mary L. Northway

Institute of Child Study
University of Toronto

Social Psychology. By Solomon E. Asch. New York: Prentice-Hall, Inc., 1952. Pp. xvi, 646. $5.50.

Here is a book that has almost everything against it. Its organization baffles the reader until he discovers that the author's original title (presumably rejected by the publishers) was "Chapters in Social Psychology." As in any series of papers, the same topic appears more than once, there is some repetition, and lines of argument are not brought together in general conclusions. Chapter headings are unfamiliar and, since the index is even more defective than usual, it is hard to guess where a given topic will be found. There is no attempt to be up to date; references are scanty and include little recent work. Worse still, the content is largely intellectual discussion, often so subtle and involved as to tax even the graduate. There are no easily grasped statements to copy out, no neat chapter summaries, no convenient tabulations of points. Indeed, there is almost nothing which the diligent student can memorize for examination—a fatal defect in a textbook. And, worst of all, the author's whole treatment is flatly out of step with the contemporary trends in our discipline. In truth, almost everything is against his book.

But not everything. Those who take the trouble to read it will discover that Asch has written a remarkable book, the fruit of long years, not only of experimentation (which any Ph.D. can do) but of hard, patient thinking, which almost no one seems to have time for. It is a book full of new and profound insights; a courageous book, which never shirks a difficult problem. Above all, it is a book of social *psychology,* which deals with social experience and behaviour in psychological terms, not in terms of physiological processes or cultural pressures.

Asch's biases are never concealed; he is a Gestaltist, a phenomenologist, who regards man's ability to understand situations and relationships as more important for social life than his primitive drives or his suscepti-

bility to simple conditioning. He looks at our scientific psychologies and concludes that a real social psychology can not be built on them. We need something different; a psychology based on the experience and behaviour of normal human beings, one which starts from the fact that psychological processes have the "extraordinary property" of providing us with *knowledge* of things, persons, and relations.

This emphasis on cognition is enough to repel the behaviourist, but Asch goes farther. He is bold enough to believe that man's millenia of social life have not left him unaltered, that social experience works a real and irreversible transformation in our psychological nature. We are not just animals who have learned new tricks, but creatures with unique and distinctive properties; in particular, the ability to make our acts socially relevant through direct perception of the motives, intentions, and feelings of others. In ironic contrast with most social psychologists, Asch takes socialization seriously; he believes that it actually changes man.

This central argument is persuasively presented in Parts III and IV, after a hundred pages devoted to criticism of biological and Freudian theories, and to a structural interpretation of motivation and learning. Later chapters deal with related topics, and several of the author's suggestive experiments on the perception of persons and the effects of social pressure are described.

Asch is too aware of the vastness of our ignorance to be dogmatic, and his arguments are often inconclusive, decision being left to the reader and the future. Not since Cooley have we had an author so modest, so tentative, so reluctant to force his views on others. This refusal ever to drive a point home, coupled with his ignoring of "student appeal" and his casual organization, may explain why Asch's book, published in 1952, is still so little known. Its real importance, however, should gradually be recognized. Some, unaccustomed to hard thinking, will find it impossible to read; others, set on applying the simplest formulations to all behaviour, will reject it as irrelevant. But in two small undergraduate courses in which it is being used a few of the brightest students respond as to a new vision, for Asch illumines great areas of social psychology which are simply ignored by other writers. This reviewer had the book for two years before he read chapters 4, 5, and 6 with the attention they require. Then, with some embarrassment, he was forced to confess that, after 25 years of teaching social psychology, he had just discovered what its subject-matter ought to be.

J. D. Ketchum

University of Toronto